



COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

FACULTAD DE ADMINISTRACIÓN DE EMPRESAS

**ANÁLISIS DE LA RED DE VIGILANCIA
DE CALIDAD DEL AIRE DEL
AYUNTAMIENTO DE MADRID
MEDIANTE MÉTODOS DE
INTERPOLACIÓN ESPACIAL**

Autor: Samuel Grau Bolívar

5º E-3 Analytics

ADE - *Business Analytics*

Tutor: Manuel Alejandro Betancourt Odio

RESUMEN

La limpieza del aire que respiramos cada día comprende una de las cuestiones de creciente interés por parte de la comunidad internacional, especialmente entre los defensores de aspectos medioambientales. Las cifras recogidas por entes independientes, como la ORGANIZACIÓN MUNDIAL DE LA SALUD, evidencian la crisis de calidad del aire a la cual nos encontramos sometidos desde hace décadas. Con cientos de miles de muertes prematuras al año por causas evitables, como las derivadas de exposiciones a niveles excesivos de contaminantes, resulta importante evaluar los sistemas de medición y reporte implementados en nuestra localidad y en escalas internacionales de forma que se busque entender el correcto funcionamiento de estas. Los métodos de interpolación espacial, como el kriging ordinario y la ponderación de la distancia inversa, desarrollados por la disciplina de la geoestadística nos permiten colmar las lagunas estadísticas posibles que existan en los sistemas de vigilancia implementados a todo nivel administrativo. En este sentido, se realiza una evaluación del sistema de vigilancia de calidad del aire implementado en el AYUNTAMIENTO DE MADRID, en relación con las concentraciones de dióxido de nitrógeno (NO₂) registradas. Así mismo, se procede a una comparativa de los métodos de interpolación espacial mencionados, desde un punto de vista visual y numérico sobre su grado certeza en las estimaciones realizadas. Por otro lado, se propone un nuevo diseño de la red, en términos de la ubicación de las estaciones de monitorización, de forma que se minimiza la varianza de cada estimación computada en todo el territorio del AYUNTAMIENTO.

Palabras Clave: Contaminación atmosférica, interpolación espacial, optimización, geoestadística.

ABSTRACT

The cleanliness of the air we breathe every day is one of the issues of growing interest to the international community, especially among environmental advocates. Figures collected by independent bodies, such as the WORLD HEALTH ORGANIZATION, highlight the air quality crisis we have been facing for decades. With hundreds of thousands of premature deaths per year due to preventable causes, such as those derived from exposure to excessive levels of pollutants, it is important to evaluate the measurement and reporting systems implemented in our locality and on international scales in order to understand their correct functioning. Spatial interpolation methods, such as ordinary kriging and inverse distance weighting, developed by the discipline of geostatistics allow us to fill the possible statistical gaps that exist in the monitoring systems implemented at all administrative levels. In this sense, an evaluation of the air quality monitoring system implemented in the Madrid CITY COUNCIL is carried out, in relation to the concentrations of nitrogen dioxide (NO₂) recorded. Likewise, a comparison of the mentioned spatial interpolation methods is made, from a visual and numerical point of view on their degree of certainty in the estimates made. On the other hand, a new design of the network is proposed, in terms of the location of the monitoring stations, in order to minimize the variance of each estimate computed in the whole territory of the CITY COUNCIL.

Keywords: Air pollution, spatial interpolation, optimization, geostatistics.

ÍNDICE

ABREVIATURAS	5
CAPITULO I. INTRODUCCIÓN	6
1. JUSTIFICACIÓN DEL PROYECTO.....	6
2. OBJETIVOS GENERALES Y ESPECÍFICOS	9
CAPITULO II. MARCO TEÓRICO	11
1. ACTUALIDAD DEL TEMA	11
1.1. Análisis de contaminantes principales. Fuentes y efectos sobre la salud:	11
1.2. Regulación sobre calidad del aire y sistemas de vigilancia nacionales.....	14
1.3. Sistemas de vigilancia internacionales y de iniciativa privada.....	19
2. REVISIÓN BIBLIOGRÁFICA.....	24
CAPÍTULO III. ANÁLISIS DE LA RED DE VIGILANCIA DEL AYUNTAMIENTO DE MADRID.....	30
1. ESTRUCTURA DE LOS DATOS. ANÁLISIS EXPLORATORIO.....	30
2. VISUALIZACIONES GEORREFERENCIADAS EXPLORATORIAS	36
3. ANÁLISIS DE MODELOS DE INTERPOLACIÓN ESPACIAL	42
3.1. Ponderación de la Distancia Inversa (IDW)	42
3.2. Método Probabilístico Kriging.....	49
CAPÍTULO IV. DISTRIBUCIÓN ÓPTIMA DE LAS ESTACIONES DE MONITORIZACIÓN DE LA RED DE VIGILANCIA DEL AYUNTAMIENTO DE MADRID.	55
CAPÍTULO V. CONCLUSIONES.....	61
BIBLIOGRAFÍA	62

ABREVIATURAS

OMS:	ORGANIZACIÓN MUNDIAL DE LA SALUD
ICA:	ÍNDICE NACIONAL DE CALIDAD DEL AIRE
CAMS:	SISTEMA DE MONITORIZACIÓN DE LA ATMÓSFERA DEL SISTEMA COPERNICUS-CAMS
BAQI:	ÍNDICE DE CALIDAD DEL AIRE DE BREEZOMETER.
KNMI:	REAL INSTITUTO METEOROLÓGICO DE HOLANDA
IDW:	PONDERACIÓN DE LA DISTANCIA INVERSA
RMSE:	RAÍZ DEL ERROR CUADRÁTICO MEDIO
SSA:	ALGORITMO DE RECOCIDO SIMULADO ESPACIAL

CAPITULO I. INTRODUCCIÓN

1. JUSTIFICACIÓN DEL PROYECTO

La calidad del aire se ha venido dando por sentado a lo largo de nuestra historia hasta décadas muy recientes. La limpieza del mismo comprende la organización y correcto funcionamiento de un sinnúmero de factores ambientales, así como de la colaboración, conciencia y regulación de los seres humanos que disfrutamos de él. Prueba del apenas reciente crecimiento en importancia de la calidad del aire y ambiente en general se constata en la RESOLUCIÓN A/RES/76/300 de 28 de julio de 2022 aprobada por la ASAMBLEA GENERAL de las NACIONES UNIDAS. En ésta, los Estados miembros y firmantes de la misma *“reconocen el derecho a un medio ambiente limpio, saludable y sostenible como un derecho humano”* (2022, pg.3). A pesar de no ser un texto normativo directamente vinculante con el Derecho interno de los Estados miembros de la organización, si comprende una directriz de comportamiento hacia los mismos y se erige como una máxima prioridad en el tratamiento de las condiciones humanitarias a nivel mundial.

Como derecho inalienable de toda persona, el respirar aire limpio se ha visto en peligro desde hace muchos años ya. Para el territorio nacional español, la tendencia de los principales contaminantes ha mostrado líneas convergentes hacia niveles cada vez más bajos en cada uno de ellos. No obstante, algunos de estos se han venido registrado en niveles inaceptables desde hace muchos años, como se demuestra en el informe de TENDENCIAS DE LA CALIDAD DEL AIRE EN ESPAÑA 2001-2021, elaborado por la SUBDIRECCIÓN GENERAL DE AIRE LIMPIO Y SOSTENIBILIDAD INDUSTRIAL del MINISTERIO PARA LA TRANSICIÓN ECOLÓGICA Y EL RETO DEMOGRÁFICO. Dentro del mismo, dos gráficos recalcan por sus cifras y se pueden ofrecer como ejemplos de la variabilidad a la que la calidad de nuestro aire se ha visto sujeta por dos décadas. El primero de ellos, expuesto en fig.1 se corresponde con el nivel de dióxido de nitrógeno medido a lo largo de los años 2001-2021, incluyendo su representación mediante un diagrama de caja y bigote y contrastado con el VALOR LÍMITE (VL), valor que corresponde al fijado en base a conocimientos técnicos y científicos *“con el fin de evitar, prevenir o reducir los efectos nocivos para la salud humana”* (SUBDIRECCIÓN GENERAL DE AIRE LIMPIO Y SOSTENIBILIDAD INDUSTRIAL, 2023, pg. 5). De forma similar, el segundo de estos gráficos se expone debajo (fig.2) y detalla la información relevante a las partículas finas de menos de 10 micrómetros de diámetro (PM 10) en el mismo formato que el anterior.

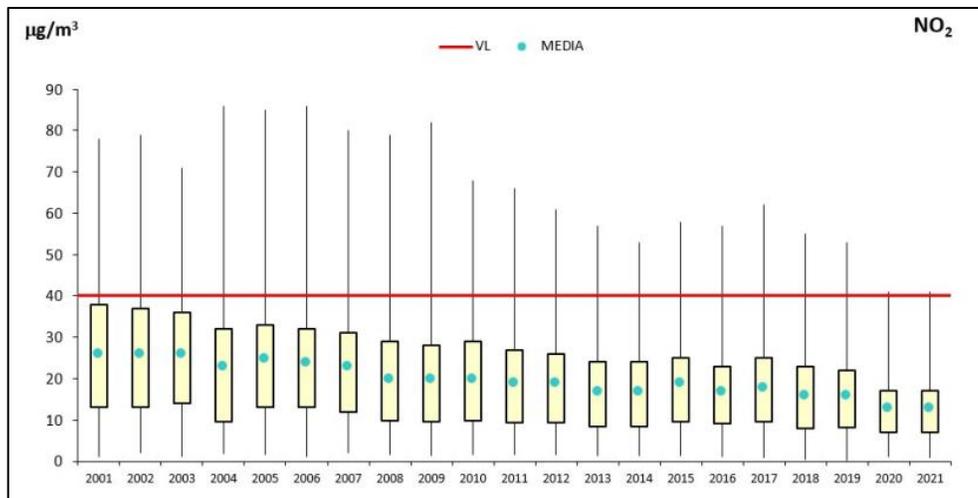


Fig. 1 Diagrama de caja y bigotes de las medias anuales de NO₂ 2001-2021 (Subdirección General de Aire Limpio y Sostenibilidad Industrial, 2023, pg. 10).

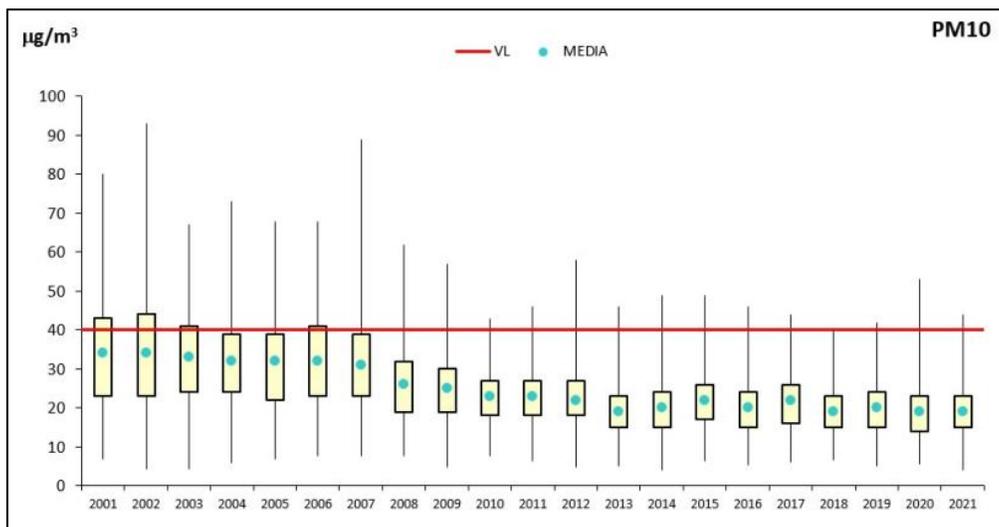


Fig. 2 Diagrama de caja y bigotes de las medias anuales de PM₁₀ 2001-2021 (Subdirección General de Aire Limpio y Sostenibilidad Industrial, 2023, pg. 18).

A pesar de observar una generalidad de niveles registrados por debajo del valor límite recomendado, la realidad también afirma que llevamos varios años expuestos en múltiples ocasiones a niveles extremadamente altos de contaminantes como éstos. Así mismo, podemos apreciar la variabilidad a las que nos veíamos sometidos sobre todo al comienzo del estudio del análisis. Todo ello cobra especial relevancia al entender que encontrarse expuesto a altos niveles de contaminación “puede aumentar el riesgo de infecciones respiratorias, enfermedades cardíacas, accidentes cerebrovasculares y cáncer de pulmón” (ORGANIZACIÓN PANAMERICANA DE LA SALUD, Impactos de la Contaminación del Aire Ambiental Exterior en la Salud, párrafo. 2).

En este mismo sentido, la ORGANIZACIÓN MUNDIAL DE LA SALUD (OMS, en adelante) ha estimado “que los efectos combinados de la contaminación del aire ambiental y la del aire doméstico se asocian a 6,7 millones de muertes prematuras cada año” (2022, Datos y cifras, párrafo 1). Con la finalidad de eliminar este extremadamente elevado número de muertes evitables debidas a la contaminación del aire, la OMS ha venido estableciendo baremos de contaminación máxima aceptable por contaminante desde el año 2000. La última actualización de esta corresponde a la elaborada en 2021 en sus DIRECTRICES MUNDIALES DE LA OMS SOBRE LA CALIDAD DEL AIRE, la cual eleva este baremo de niveles de calidad del aire a una exigencia mayor basándose en la evidencia recopilada desde la elaboración del informe publicado en el 2000. Aún más, se establecen niveles recomendados en función de las metas intermedias que conllevarían a una verdadera consecución del objetivo en reducir estas muertes evitables.

Contaminante	Tiempo promedio	Meta intermedia				Nivel de las directrices sobre la calidad del aire
		1	2	3	4	
MP _{2,5} , µg/m ³	Anual	35	25	15	10	5
	24 horas ^a	75	50	37,5	25	15
MP ₁₀ , µg/m ³	Anual	70	50	30	20	15
	24 horas ^a	150	100	75	50	45
O ₃ , µg/m ³	Temporada alta ^b	100	70	–	–	60
	8 horas ^a	160	120	–	–	100
NO ₂ , µg/m ³	Anual	40	30	20	–	10
	24 horas ^a	120	50	–	–	25
SO ₂ , µg/m ³	24 horas ^a	125	50	–	–	40
CO, mg/m ³	24 horas ^a	7	–	–	–	4

^a Percentil 99 (es decir, 3-4 días de superación por año).
^b Promedio de las concentraciones máximas diarias de O₃ (medias octohorarias) en los seis meses consecutivos con la concentración media móvil de O₃ más alta.

Fig. 3 Niveles recomendados de las directrices sobre la calidad del aire y metas intermedias (OMS, 2021, pg. 5).

En directa contravención de los estándares mínimos de limpieza del aire impuestos por la OMS, todos los españoles y millones de europeos han visto perjudicada su salud por los niveles de partículas y contaminantes que encontramos cada día en nuestros aires. En evidencia de ello, la OMS registró en el 2019 que “el 99% de la población mundial vivía en lugares donde no se respetaban las Directrices de la OMS sobre la Calidad del Aire” (OMS, 2022, Datos y cifras, párrafo 1).

En España, particularmente, las partículas en suspensión (PM10 y PM2,5), el dióxido de nitrógeno (NO2) y el ozono troposférico (O3) resultan los más problemáticos. Dichos contaminantes han afectado la vida de uno de cada diez españoles en el 2021, así como un 24,2% de toda la vegetación del territorio español, tomando en cuenta la normativa actual. No obstante, de cara al futuro, se podría considerar como contaminado un 80% del territorio nacional y su inevitable consecuencia en términos de salud para su población. (ECOLOGISTAS EN ACCIÓN, 2022)

Se computaron 400.000 muertes prematuras en la mayoría de Europa por la calidad y contaminación del aire en 2019. Para España, esto logró representar una cantidad de 30.000 vidas perdidas por respirar cada día aire contaminado, en su gran mayoría por partículas en suspensión (PM2,5) y en menor medida por NO2 y por O3, duplicando los fallecimientos anuales evitables que se estimaban hace dos décadas. Así mismo, esta pérdida de vida prematura junto con la merma en productividad de los trabajadores ya sea por pérdida de días trabajados o menores funcionalidades por perjuicios a la salud, han llegado a suponer unos 38 millones de euros anuales en 2013, siendo equivalentes a un 3,5% del PIB del estado español. Todo ello se puede ver aumentado hoy en día, además de tomar en cuenta la falta de estimación del daño a la vegetación para el cálculo económico antes mencionado. (ECOLOGISTAS EN ACCIÓN, 2022)

2. OBJETIVOS GENERALES Y ESPECÍFICOS

Por todo ello, este Trabajo de Fin de Grado busca realizar un estudio sobre las maneras en las que estas mediciones de partículas y contaminantes se han venido llevado a cabo.

En primer lugar, se elaborará una extensa revisión del estado actual de la cuestión en el ámbito local (AYUNTAMIENTO DE MADRID), en el ámbito regional (COMUNIDAD AUTÓNOMA DE MADRID), en el ámbito nacional (Estado español) y en el ámbito internacional (UNIÓN EUROPEA e iniciativas privadas globales). Se trata, por tanto, de una recopilación de los distintos sistemas que han venido implementándose a estos niveles administrativos para la monitorización y reporte sobre la calidad del aire en los distintos territorios que los comprenden, así como de la regulación vigente que trata dicha cuestión con el ánimo de salvaguardar el derecho a un aire limpio.

De forma similar, se buscará indagar en métodos de estimación de niveles de contaminación del aire para puntos específicos a partir de mediciones en estaciones meteorológicas fijas y recolectados con sus instrumentos de medición. De forma más específica, se buscará desarrollar

un sistema de evaluación de las distintas maneras en las que se puedan estimar niveles de contaminación a través de datos y puntos conocidos. Se hará énfasis en formas de interpolación espacial, dentro de la disciplina de la geoestadística, como métodos de cómputo de valores desconocidos a partir de valores observados en distintos puntos. Dentro de éstos, se verán implementados métodos de interpolación espacial determinísticos, como lo puede ser el método de la ponderación de la distancia inversa, así como métodos de interpolación espacial probabilísticos, como lo es el método Kriging Ordinario.

Dicho sistema de evaluación buscará comparar cada el desempeño de ambos métodos en estimación de concentraciones de contaminantes para puntos desconocidos, contrastando los mismo con los valores observados y disponibles en el Trabajo. Para su implementación, se utilizará una serie de paquetes y librerías del lenguaje de programación R y se buscará su automatización mediante el mismo. El código implementado para la elaboración de todo el contenido del Trabajo se encuentra en el siguiente enlace: [2 Items](#).

Así mismo, se buscará realizar dicho análisis comparativo de métodos de estimación espacial para las distintas concentraciones de dióxido de nitrógeno (NO₂), circunscribiéndonos a la unidad territorial del AYUNTAMIENTO DE MADRID. Como podremos ver más adelante, resulta de especial interés conocer las concentraciones de dicho contaminante por sus particulares efectos perjudiciales para la salud humana, vegetal y animal, así como por sus características precursoras y determinantes de niveles de concentración de contaminantes ulteriores. De forma similar, la región administrativa-territorial del AYUNTAMIENTO DE MADRID abarca la mayor densidad poblacional de las regiones estudiadas, con la más de la mitad de los habitantes de toda provincia (INE, 2022).

Finalmente, se busca estudiar la idoneidad del diseño actual de la red de vigilancia de la calidad del aire en el AYUNTAMIENTO DE MADRID, para el óptimo registro de concentraciones de contaminantes como el dióxido de nitrógeno, a lo largo de todo su territorio. Se evaluarán las regiones más precisas, en cuanto a su cobertura y precisión, y se intentará ofrecer una configuración que minimice los posibles errores de estimación computados con los métodos de interpolación espacial mencionados previamente.

CAPITULO II. MARCO TEÓRICO

1. ACTUALIDAD DEL TEMA

1.1. Análisis de contaminantes principales. Fuentes y efectos sobre la salud:

Como ya hemos podido comentar brevemente, la OMS se ha beneficiado de la recopilación de datos a lo largo de las últimas dos décadas para el estudio y determinación de los efectos de los contaminantes que encontramos en nuestro aire sobre la salud todos los seres humanos, especialmente en términos de causalidad. De forma similar, y como también hemos mencionado previamente, los contaminantes más dañinos en suelo español se reducen a algunos de los contaminantes más típicos (PM_{2,5}; PM₁₀; O₃; y NO₂). Ahora bien, sería interesante entender lo que la OMS y otras instituciones han venido concluyendo acerca de los mismo en los últimos años.

Dentro de los contaminantes más nocivos para la salud de todos los seres humanos se encuentran las materias particuladas en suspensión de diversos diámetros. Las más comunes de éstas son las de 2,5 y 10 micrómetros de diámetro; PM_{2,5} y PM₁₀, respectivamente. Tanto las primeras como las segundas comparten muchas de sus fuentes de producción y emisión al aire que respiramos. Dentro de éstas encontramos fuentes tanto naturales, como lo pueden ser tormentas de arenas o incendios forestales, así como aquellas procedentes de actividad humana, *“tráfico, especialmente vehículos diésel, incineradoras, calefacciones de carbón, minería, procesos industriales, etc.”* (COMUNIDAD DE MADRID, Partículas (PM_{2,5} y PM₁₀), párrafo 1). No obstante, resulta importante mencionar que sí existe una diferenciación aparente en cuanto a la intensidad de las emisiones de las diversas fuentes mencionadas y la concentración de cada tipo de materia particulada, afectando así la calidad del aire en su proximidad de una forma determinada. En efecto, las partículas PM_{2,5} proceden fundamentalmente de las fuentes de actividad humana y, por tanto, serán más prevalentes en aquellas zonas emisoras caracterizadas por su industrialización y su proximidad a redes de tráfico intensas. Por otra parte, las partículas de 10 micrómetros de diámetro, por lo general, son provenientes de fuentes naturales y, en particular en el territorio español, se pueden ver ejemplificadas por la subida de polvo o arena del desierto de Sahara cada año (COMUNIDAD DE MADRID, Partículas (PM_{2,5} y PM₁₀), párrafo 1).

En cuanto al ozono, sabemos que se trata de un compuesto encontrado de manera natural en la atmósfera. Como la tan importante capa de ozono, ubicada aproximadamente a 20 km de la superficie terrestre, este compuesto tiene la función de proteger a todos los seres vivos de las

radiaciones ultravioleta emitidas en todo nuestro sistema solar (GENERALITAT DE CATALUNYA, El ozono troposférico). No obstante, es importante mencionar, que además de encontrarse este compuesto depositado en dicha capa protectora de nuestro planeta, denominado ozono estratosférico, también está presente en la sección de la atmósfera más cercana a la superficie, la tropósfera, y por ello se le denomina a este ozono troposférico. En esta ubicación, su función es totalmente opuesta a la de protección de nuestras vidas, ya que constituye un contaminante de alta peligrosidad. La complejidad de este oxidante fotoquímico radica en que se trata de un contaminante secundario, siendo este uno que no se ve emitido directamente a la atmósfera por una fuente como lo pueden ser las materias particuladas descritas previamente, sino que éste se ve *“formado a partir de reacciones fotoquímicas (activadas por la luz solar) entre contaminantes primarios. Concretamente, se forma ozono cuando coexisten los óxidos de nitrógeno (NO_x), los compuestos orgánicos volátiles (COVs) y una radiación solar intensa a lo largo de un periodo de tiempo lo bastante largo (un mínimo de diversas horas)”* (GENERALITAT DE CATALUNYA, El ozono troposférico: un contaminante secundario). Por ello, podemos decir que este contaminante se forma en presencia de otros emitidos principalmente por fuentes de actividad humana, como lo son los óxidos de nitrógeno, a través de las reacciones químicas entre la luz solar y las emisiones de automóviles, complejos industriales, procesos de minería, etc.

Por último, pero tal vez con la mayor importancia, resulta indispensable comentar los óxidos de nitrógeno como contaminantes. La atmósfera se ve contaminada, principalmente, por dos compuestos abarcados en esta categoría; el óxido nítrico (NO) y el dióxido de nitrógeno (NO₂). La fuente principal de ambos de estos compuestos se reduce a los procesos de combustión de todo vehículo o dispositivo que se encuentre funcionando mediante un motor a combustión. Se considera que específicamente aquellos motores que utilicen el diésel como precursor de energía producen mucha mayor cantidad de estos óxidos de nitrógeno (COMUNIDAD DE MADRID, Dióxido de Nitrógeno (NO₂), párrafo 1). No obstante, es importante entender que el proceso general de formación de estos óxidos inicia con la mezcla de oxígeno y nitrógeno a temperaturas elevadas en el proceso de combustión de los motores, conllevando a la formación de óxido nítrico (NO) en primer lugar. Una vez este contaminante primario es emitido, gran parte del NO₂ de la atmósfera se formará como consecuencia de las reacciones químicas que sufre el NO con los oxidantes ya presentes en la misma, entre otros, el ozono troposférico (GOBIERNO DE LA RIOJA, Óxidos de Nitrógeno, párrafo 1).

Sabiendo las principales fuentes de emisión de cada uno de estos contaminantes, tanto primarios como secundarios, es importante entender cuáles son los efectos que tiene cada uno de ellos sobre la salud tras la exposición de los seres humanos a los mismos. En primer lugar, resulta claro que todos ellos producen una disminución notable de la función pulmonar, así como un incremento en la probabilidad de sufrir de enfermedades cardíacas y accidentes cardiovasculares (ORGANIZACIÓN PANAMERICANA DE LA SALUD, Impactos de la Contaminación del Aire Ambiental Exterior en la Salud, párrafo 2). Ahora bien, se ha venido observado una importante correlación entre los reportes de ataques de asma con niveles altos de ozono troposférico exposiciones prolongadas al mismo. Además de ello, se han un aumento en el número de urgencias atendidas por problemas respiratorios a medida que incrementa la concentración de este contaminante (O₃) (COMUNIDAD DE MADRID, Ozono (O₃), párrafo 4). Así mismo, este contaminante se ve ligado a un incremento en la mortalidad de todo tipo y mortalidad respiratoria en su exposición a largo plazo.

En cuanto a las materias particuladas (PM_{2,5} y PM₁₀), los efectos sobre la salud se ven intensificados. Especialmente con las partículas de 2,5 micrómetros de diámetro, la exposición a las mismas puede resultar perjudicial ya que pueden ingresar fácilmente por las vías respiratorias y, por su tamaño, *“pueden penetrar hasta las partes más profundas del pulmón y pasar a la sangre”* (COMUNIDAD DE MADRID, Partículas (PM_{2,5} y PM₁₀), párrafo 2). Este acceso casi irrestricto del contaminante al torrente sanguíneo conlleva a la mayor parte de las incidencias de enfermedades cardiovasculares y cerebrovasculares registradas por exposición a un aire que no cumpla con los baremos de limpieza de la OMS. Por todo ello, se ha concluido que la exposición a largo plazo, en particular de materias particuladas de 2,5 y 10 micrómetros de diámetro incrementan la mortalidad por toda causa, por causa cardiovascular (incluidas las cerebrovasculares) y por causas pulmonares, además de mostrar clara causalidad de cáncer de pulmón (OMS, 2021).

Finalmente, los óxidos de nitrógeno también resaltan por sus efectos sobre la salud. Particularmente, el dióxido de nitrógeno (NO₂) se caracteriza por ser de color pardo-rojizo y tener un olor tóxico y asfixiante. Ello resulta, principalmente, en la irritación de los accesos a las vías respiratorias (nariz y boca) así como el del resto de partes que conducen a la mucosa como lo pueden ser los ojos. Uno de los síntomas más comunes de su exposición es el de tos y dificultades para respirar. De forma similar al resto de contaminantes mencionados, la exposición a largo plazo

de este incrementa el riesgo de mortalidad por toda causa y el riesgo de mortalidad por causas o complicaciones respiratorias en los seres humanos (OMS, 2021). Por otra parte, sin embargo, el dióxido de nitrógeno se muestra perjudicial en cualquier entorno además del de la vida humana. Su presencia en nuestro aire determina la acumulación y deposición en capas superficiales de la atmósfera, las cuales originan fenómenos atmosféricos de extremo peligro como los son las lluvias ácidas, contribuyendo así a la “*destrucción de la vegetación y la eutrofización de lagos y océanos*” (FERNÁNDEZ RIVAS, 2015 p. 2).

1.2. Regulación sobre calidad del aire y sistemas de vigilancia nacionales

Por todo ello, y como se expuso brevemente al inicio del trabajo, la OMS ha venido estableciendo determinados tramos de niveles de contaminación recomendado para cada tipo de contaminante con la finalidad de fomentar un aire más limpio y, al mismo tiempo, establecer un límite numérico medible que asegure la preservación de la salud de todos los seres vivos. Como tuvimos oportunidad de establecer previamente, estos niveles recomendados han sido determinados desde el año 2000 por la OMS. Actualmente, prevalecen los límites establecidos en las DIRECTRICES MUNDIALES DE LA OMS SOBRE LA CALIDAD DELAIRE, en su versión del 2021, y vistos en este trabajo en la figura 3.

Por otra parte, la gran mayoría de los países europeos regulan las distintas consideraciones sobre la calidad del aire sobre su territorio nacional. En el caso de España, estas consideraciones se ven consagradas en el REAL DECRETO 102/2011, DE 28 DE ENERO, RELATIVO A LA MEJORA DE LA CALIDAD DELAIRE. Dentro de este texto normativo se recogen las aspiraciones del gobierno español en la consecución de objetivos como mantener un aire limpio, medir cada uno de los contaminantes, asignar determinados organismos públicos las distintas competencias para evaluar y gestionar la calidad, entre otras cosas.

En este mismo sentido, en el 2019 se aprueba el **ÍNDICE NACIONAL DE CALIDAD DEL AIRE (ICA)** en España. Se trata de un sistema de medición y reporte que ayuda a *“representar la calidad del aire a nivel nacional de una manera fácilmente entendible por los ciudadanos contribuyendo al acceso del público a dicha información ambiental de una manera clara”* (Orden TEC/351/2019, Preámbulo, párrafo 8). De manera más precisa, el ICA comprende una amplia red de estaciones de monitorización que registra en cada una de sus ubicaciones el nivel de los siguientes contaminantes: Partículas en suspensión (PM10), Partículas en suspensión (PM2,5), Ozono troposférico (O3), Dióxido de nitrógeno (NO2) y Dióxido de azufre (SO2). De este registro horario de cada contaminante, se nutre una base de datos que permite evaluar la calidad del aire en dichos puntos, establecer visualizaciones a través de un visor en tiempo real de los datos en cada estación, así como la evolución histórica de la calidad del aire en cada punto, y que, además, permite proveer recomendaciones de actuación al aire libre para las poblaciones cercanas a cada estación respectivamente.

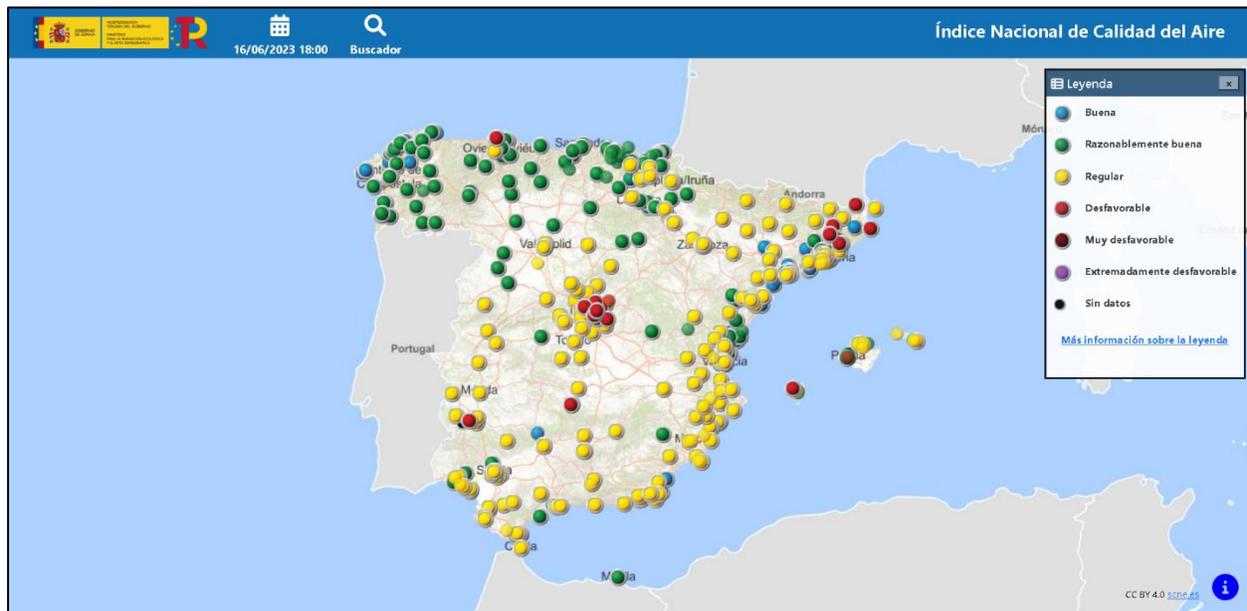


Fig. 4 Visor de calidad del aire del sistema ICA (MINISTERIO PARA LA TRANSICIÓN ECOLÓGICA Y EL RETO DEMOGRÁFICO, Índice Nacional de Calidad del Aire).

Como se puede apreciar en la figura 4, la red de estaciones de monitorización que compone el ICA abarca todo el territorio nacional. Se trata de más de 600 estaciones de medición fijas y de más de 4000 analizadores individuales. La gestión de todas ellas corresponde, respectivamente, a cada COMUNIDAD AUTÓNOMA y ENTIDAD LOCAL en donde se ubique la estación de monitorización. Estas entidades territoriales “dividen todo su territorio en zonas y aglomeraciones, basándose en criterios de homogeneidad en cuanto a tipo y concentración de contaminantes, así como de gestión administrativa de las mismas” (MINISTERIO PARA LA TRANSICIÓN ECOLÓGICA Y EL RETO DEMOGRÁFICO, Calidad del aire, párrafo 3). El sitio web del MINISTERIO que trata el ICA permite la redirección a los sitios de cada COMUNIDAD AUTÓNOMA y ENTIDAD LOCAL para la prestación de servicios similares al provisto a esta escala nacional. En este sentido, la COMUNIDAD DE MADRID, por ejemplo, ofrece un visor de capacidades muy equiparables a las del MINISTERIO sobre las 24 estaciones presentes en la provincia, sin detallar de manera extensa las ulteriores 24 estaciones que componen la red en el AYUNTAMIENTO DE MADRID, específicamente, teniendo esta última su propio portal con un visor y base de datos con capacidad potentes.

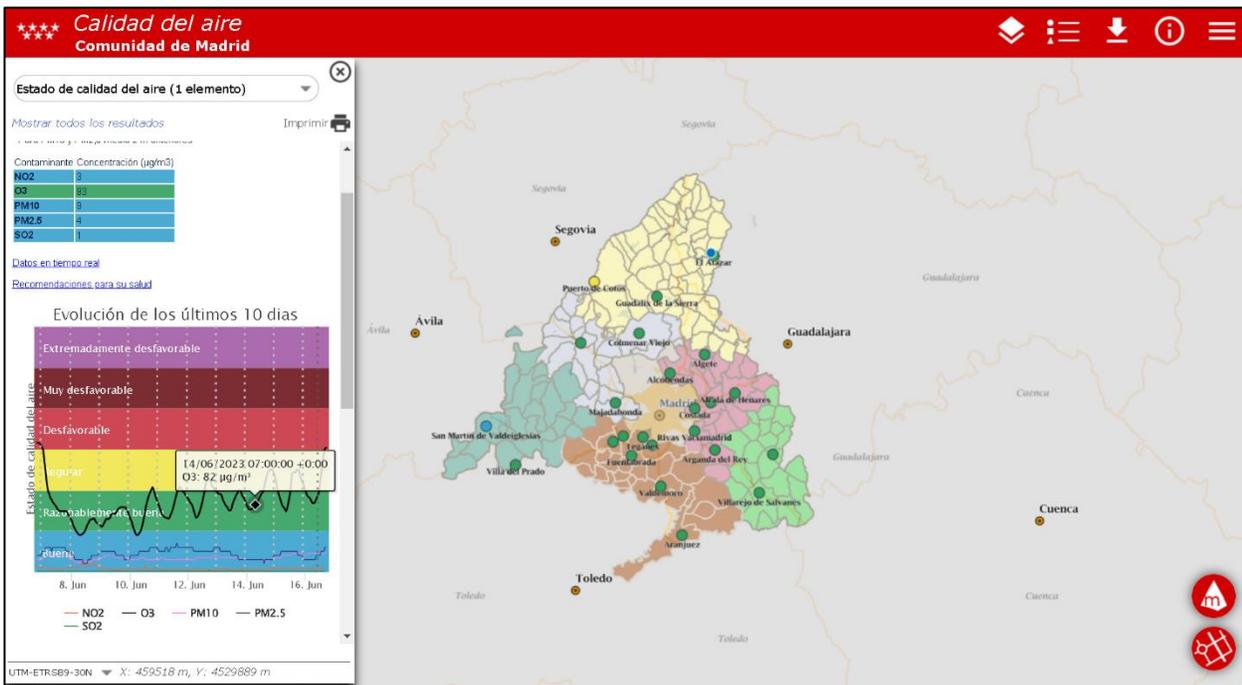


Fig. 5 Visor de Calidad del aire de la COMUNIDAD DE MADRID (COMUNIDAD DE MADRID, Calidad del aire).

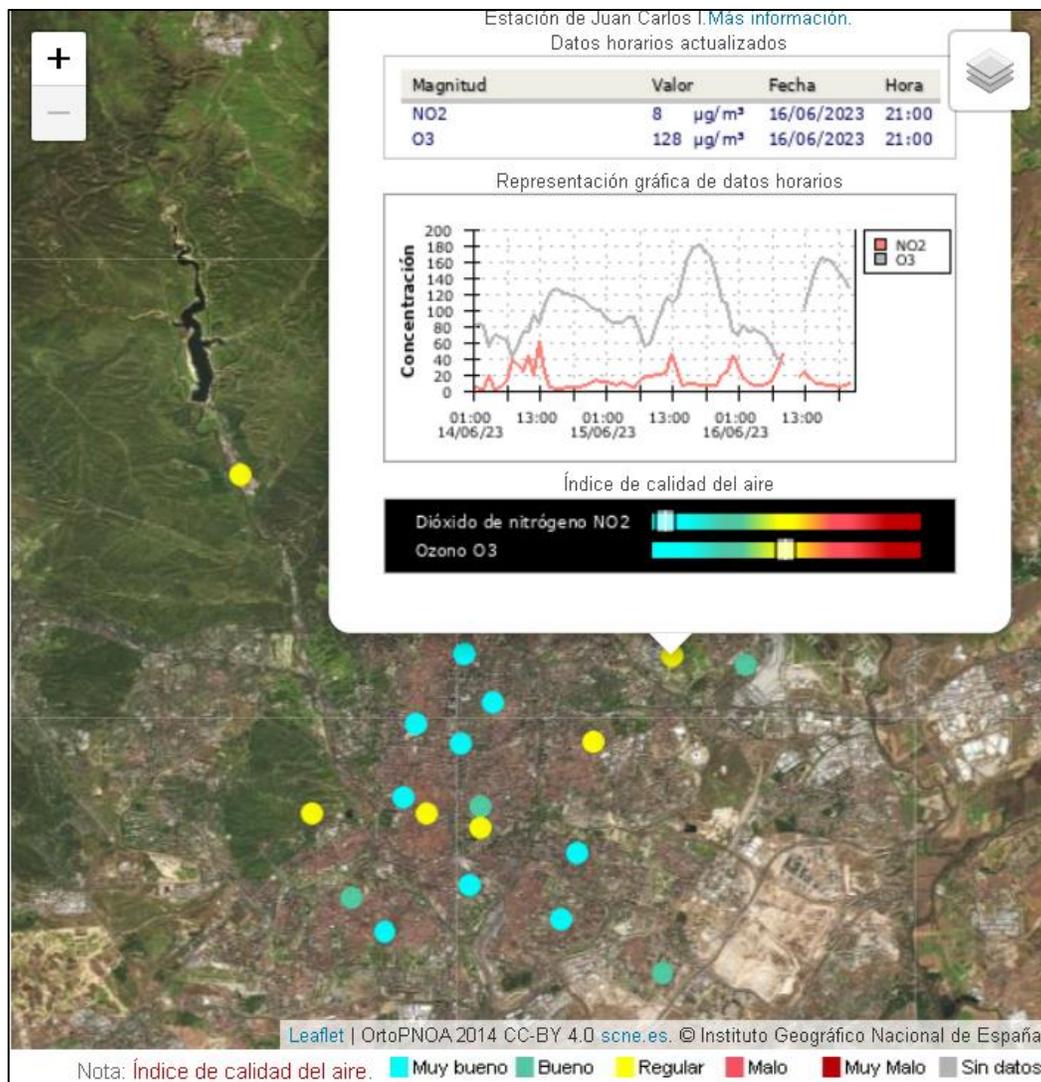


Fig. 6 Visor de calidad del aire del AYUNTAMIENTO DE MADRID (AYUNTAMIENTO DE MADRID, Mapa de la Red de Vigilancia).

Ahora bien, el **ÍNDICE NACIONAL DE CALIDAD DEL AIRE (ICA)**, como su nombre bien indica, comporta un estadístico destinado a indicar el nivel de contaminación actual del aire en cada zona y, por tanto, reflejar *“el impacto potencial de la calidad del aire sobre la salud”* (ANEXO A LA ORDEN TEC/351/2019, Metodología del cálculo del Índice Nacional de Calidad del Aire, párrafo 3). Para su cómputo se toman en cuenta los niveles registrados de los contaminantes mencionados previamente, es decir: Partículas en suspensión (PM10), Partículas en suspensión (PM2,5), Ozono troposférico (O3), Dióxido de nitrógeno (NO2) y Dióxido de azufre (SO2). Así mismo, cada estación se ve clasificada *“de acuerdo con la fuente de emisión predominante en dicha localización, de la siguiente manera: estación de tráfico, estación industrial, estación de fondo”*

(ANEXO A LA ORDEN TEC/351/2019, Metodología del cálculo del Índice Nacional de Calidad del Aire, párrafo 2). Como podremos comprobar posteriormente y, con la información que se ha expuesto previamente en este apartado, será evidente que la mayor concentración de contaminantes provenientes de fuentes de actividad humana (tráfico e industrial) esté presente en aquellas estaciones de monitorización clasificadas respectivamente. De forma similar, se ha desarrollado un sistema de bandas de índice de calidad para referenciar categóricamente la calidad del aire presente en un momento determinado en cada estación de monitorización. Su elaboración ha partido de la consideración de los riesgos relativos asociados, a corto plazo, a la exposición de los contaminantes PM_{2,5}, O₃ y NO₂, establecidos por la OMS y vistos en la figura 3 de este trabajo. El sistema de bandas del ICA se puede ver en la figura 7 a continuación.

SO ₂		PM _{2,5}		PM ₁₀		O ₃		NO ₂		CATEGORÍA DEL ÍNDICE
0	100	0	10	0	20	0	50	0	40	BUENA
101	200	11	20	21	40	51	100	41	90	RAZONABLEMENTE BUENA
201	350	21	25	41	50	101	130	91	120	REGULAR
351	500	26	50	51	100	131	240	121	230	DESFAVORABLE
501	750	51	75	101	150	241	380	231	340	MUY DESFAVORABLE
751-1250		76-800		151-1200		381-800		341-1000		EXTREMADAMENTE DESFAVORABLE

* Los valores de todos los contaminantes de la tabla están expresados en µg/m³

Fig. 7 Sistema de bandas y categorías ICA (Anexo a la Orden TEC/351/2019, Metodología del cálculo del Índice Nacional de Calidad del Aire).

De manera más precisa, el ICA se computa de acuerdo con un promedio temporal diferenciado para cada contaminante descrito en la tabla anterior. Por ejemplo, para el dióxido de nitrógeno (NO₂) y para el dióxido de azufre (SO₂) se utiliza la concentración media de la última hora registrada en la estación de monitorización. Para el ozono troposférico (O₃), el valor de referencia es la media móvil de las concentraciones de las últimas 8 horas. Finalmente, para las materias particuladas (PM_{2,5} y PM₁₀), se recurre a la media móvil de las concentraciones registradas en la estación en las últimas 24 horas. El valor final del ICA tomará el valor de la peor categoría en términos de calidad del aire, considerando todos los contaminantes que se registran en cada estación de monitorización (ANEXO A LA ORDEN TEC/351/2019, Metodología del cálculo del Índice Nacional de Calidad del Aire).

Por otra parte, resulta vital indicar que, a pesar del grado de sofisticación y de alta capacidad de cada uno de los analizadores individuales y de las estaciones de monitorización que componen la red de vigilancia de calidad del aire, existe la posibilidad de no registrar la concentración de cualquier contaminante en determinadas horas. Este fallo se puede deber a múltiples razones. De manera más frecuente, la pérdida de datos puede ocurrir por un nivel sumamente reducido de las concentraciones de cada contaminante en cada ubicación de las estaciones de monitorización. Igualmente, se puede deber a la falta de un correcto funcionamiento de los sensores o analizadores integrados en cada estación de monitorización en un determinado espacio temporal.

1.3. Sistemas de vigilancia internacionales y de iniciativa privada

Ante esta problemática de posible insuficiencia de datos, el ICA utiliza datos provenientes del SERVICIO DE VIGILANCIA ATMOSFÉRICA DE LA UNIÓN EUROPEA y, propiamente, del SISTEMA DE MONITORIZACIÓN DE LA ATMÓSFERA DEL SISTEMA COPERNICUS-CAMS (en adelante, CAMS). COPERNICUS comprende el programa de observación de la Tierra de la Unión Europea. Se encarga de monitorizar nuestro planeta y todo lo que sucede en el en términos ambientales con el fin de brindar el máximo beneficio para todos los ciudadanos de los países que conforman la Unión Europea. A pesar de ser la COMISIÓN EUROPEA la que gestiona dicho programa, se trabaja juntamente con todos los Estados miembros y con un sinnúmero de agencias estatales y europeas en la consecución de este gran objetivo. Resumidamente, *“Se utilizan enormes cantidades de datos globales procedentes de satélites y sistemas de medición terrestres, aéreos y marítimos para proporcionar información que ayude a los proveedores de servicios, las administraciones públicas*

y otras organizaciones internacionales a mejorar la calidad de vida de la ciudadanía europea” (COPERNICUS, Sobre Copernicus, párrafo 3).

Ahora bien, el sistema CAMS comprende la sección del programa COPERNICUS que busca recopilar toda la información relativa a la capa atmosférica situada sobre el espacio de la Unión. Así como todo el conjunto de medición de COPERNICUS, CAMS utiliza satélites y estaciones de monitorización *in situ* para el registro de medidas y concentraciones de contaminantes y otros compuestos orgánicos como el polen, con el objetivo de generar una imagen real de la situación actual de la atmósfera en toda esta área. Aún más, el sistema CAMS combina los datos recopilados por dichos satélites y estaciones de monitorización con modelos numéricos que ajustan la trayectoria de los gases observados en la capa atmosférica, así como para realizar predicciones sobre los valores futuros de estas concentraciones en rangos espaciales concretos de toda la Unión. Además de estar integrado por más de 40 satélites, el sistema de vigilancia global abarca miles de estaciones de monitorización situadas en la superficie terrestre y un número extenso de aeronaves, embarcaciones y globos atmosféricos, denominados campos de observación *in-situ*, que no solo se limitan a la recopilación de datos sobre contaminantes, sino también extienden su ámbito de vigilancia a factores como temperatura, condiciones nubosas, factores del viento, concentraciones de óxidos de carbono y aerosoles (CAMS, Observations).

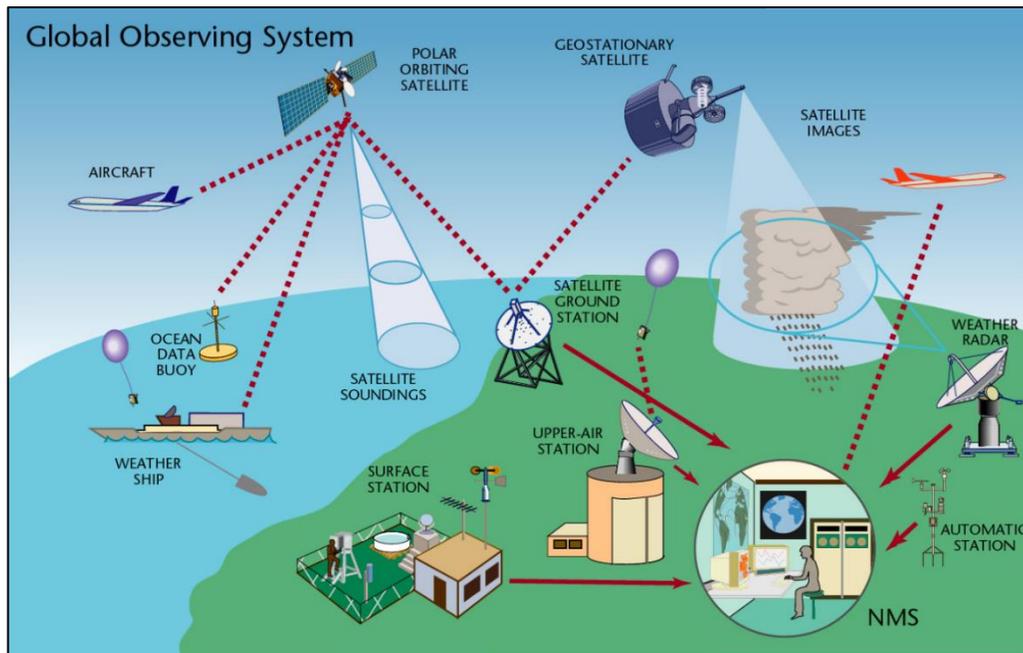


Fig. 8 Sistema de observación global CAMS (CAMS, Observations).

Predeciblemente, el portal web del sistema CAMS posee todo tipo de funcionalidades de visualización y extracción de datos de tal manera que resulta de un muy alto grado sofisticación y apto para la utilización de los mismo por parte de grupos científicos en labores de investigación. A continuación, en la figura 9, se puede observar una de las capacidades de visualización de este sistema en la representación gráfica de la calidad del aire tomando en cuenta las concentraciones diarias de materias particuladas de 10 micrómetros de diámetro (PM10) en el ámbito espacial de la Unión Europea.

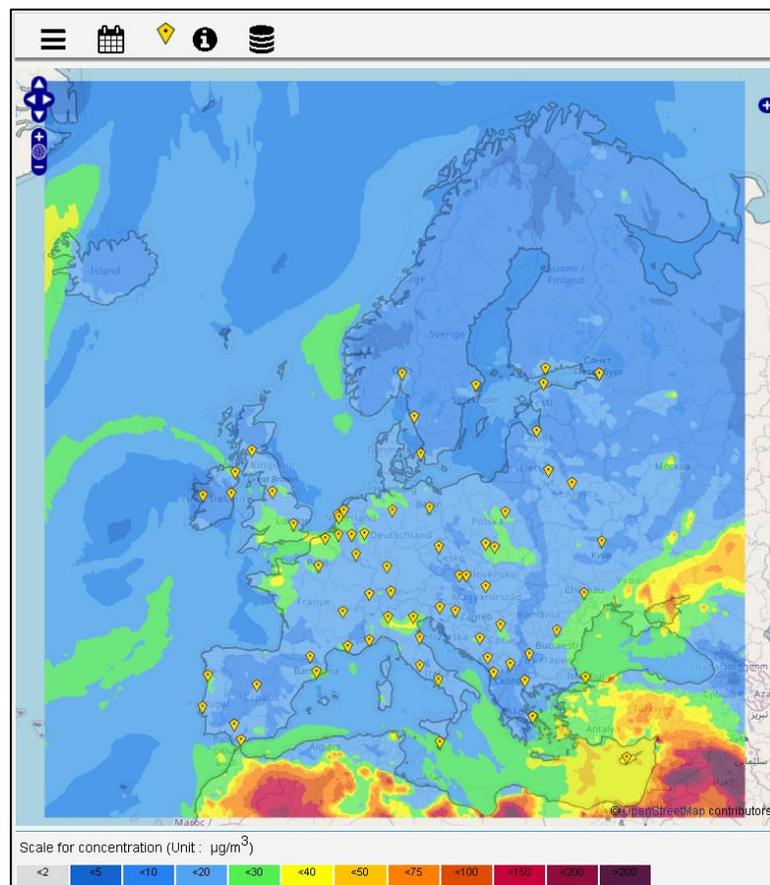


Fig. 9 Visor de calidad del aire del sistema CAMS para las concentraciones diarias de PM10 (CAMS, Air Quality in Europe).

Por otra parte, el sector privado ha desarrollado una serie aplicaciones e interfaces que permiten difundir información, así como lo hace el ICA o CAMS, a los ciudadanos de gran parte del mundo, pero con mayor agilidad. BREEZOMETER y LOBELIA muestran ser compañías pioneras en este ámbito. Por un lado, BREEZOMETER recopila información de más de 14.000 fuentes distintas

alrededor del mundo. Además de utilizar los datos satelitales del sistema COPERNICUS, utiliza información satelital de otras agencias estatales de Estados no miembros del espacio europeo. Así mismo, incorpora la información recogida por miles de estaciones de monitorización a nivel global, reportes de intensidad y flujo de tráfico, data meteorológica y también información detallada sobre la cubierta terrestre (BREEZOMETER & GOOGLE, How BreezoMeter's Air Quality Technology Works, párrafo 3). Inspirados en la búsqueda de información detallada y específica sobre la ubicación geográfica particular de cada individuo y la calidad del aire a la que se pueda encontrar expuesto, la propuesta de valor de BREEZOMETER y de sus sistemas de registro, visualización y recomendación radica en la resolución a la cual pueden aportar información a tiempo real y de forma predictiva. A diferencia de los sistemas de vigilancia de calidad del aire aquí expuestos, BREEZOMETER tiene la capacidad de reportar información sobre la calidad del aire con una granularidad de 5 metros, es decir a nivel de cada calle en una gran ciudad, por ejemplo. Utilizando la gran cantidad de datos a los que tienen acceso, se emplean técnicas de big data, machine learning e interpolación espacial para poder reporte a esta resolución.

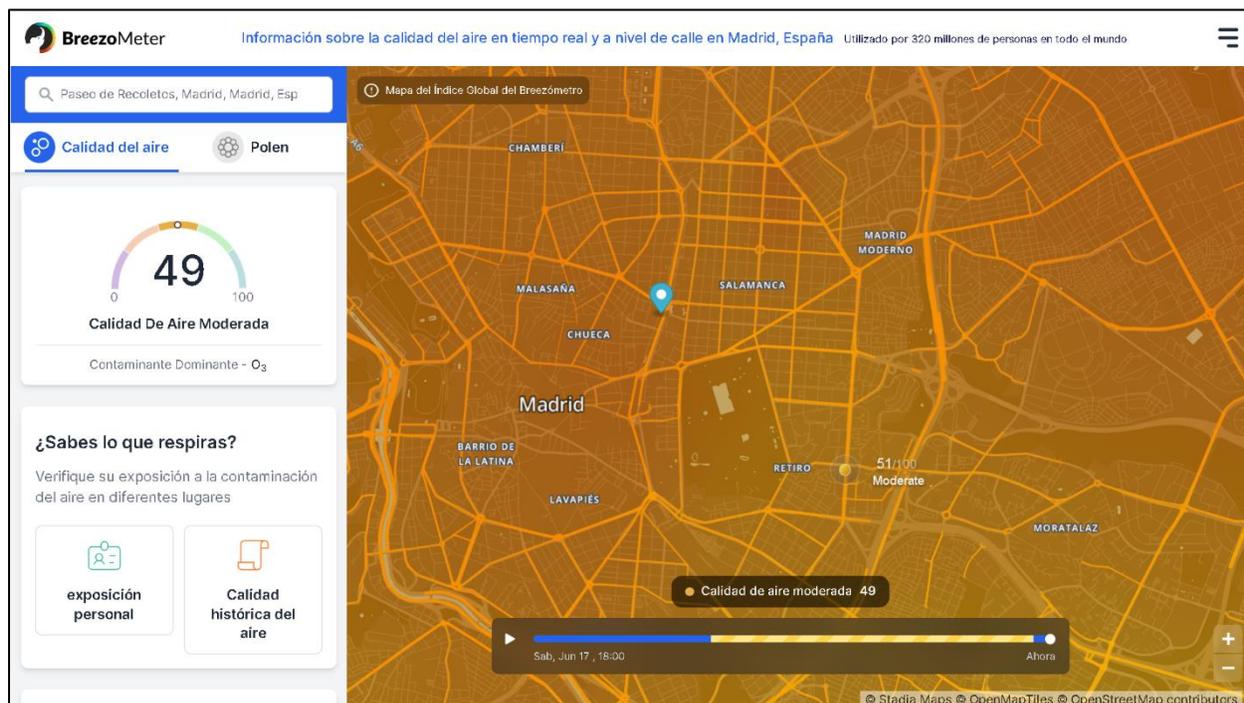


Fig. 10 Mapa de calidad del aire en el Ayuntamiento de Madrid y por calles específicas (BREEZOMETER, Air Quality Map).

Por tanto, esta revolucionaria tecnología no solo permite reportar y difundir información acerca de la calidad del aire de una manera sencilla de entender y visualizar, puesto en evidencia por estar integrada con aplicaciones móviles del clima en muchos dispositivos móviles, sino que introduce

una solución a la gran crítica que se realiza a las redes de vigilancia institucionales que se han visto aquí expuestas; el nivel de resolución o granularidad de su reporte. Por una parte, el sistema ICA dificulta conocer el nivel de contaminación en el aire que existe en cualquier otra área que no corresponda con la exacta ubicación de alguna de sus estaciones de monitorización, además de incluir ciertos problemas de pérdidas de datos incluso en estas localizaciones, como hemos mencionado previamente. Por otra, la dependencia de imágenes satelitales, como lo tiene el sistema CAMS, puede estar sujeta a obstrucciones por las condiciones nubosas de las áreas que cubren, produciendo posibles datos falsos además de resultar imposible conocer la concentración de contaminantes a una resolución como lo pueden ser cada 5 metros.

Además de todas estas cosas, el sistema integrado a BREEZOMETER monitoriza más de 12 contaminantes y componentes orgánicos de interés y su cobertura abraza más de 90 países a nivel global. Igual de interesante resulta entender que esta plataforma se ajusta una gran cantidad de índices de calidad del aire, los cuales varían sus especificaciones y métodos de cómputo dependiendo de la jurisdicción aplicable. Así mismo, introducen su propio índice de calidad del aire, BreezoMeter's Air Quality Index (BAQI), el cual se basa en investigaciones académicas y en el conjunto de otros índices globales utilizados por instituciones de la más alta fiabilidad (BREEZOMETER & GOOGLE, BreezoMeter's Universal & Local Air Quality Indexes, párrafo 4).

De forma parecida a BREEZOMETER, LOBELIA es una compañía que se encarga de utilizar *“datos de observación de la Tierra y proyecciones climáticas para evaluar vulnerabilidades, riesgos y oportunidades asociadas a los escenarios climáticos para cada caso particular y región de interés”* (LOBELIA, Nosotros, párrafo 1). Para ello, utilizan fuentes de información satelital y, principalmente, utilizan los datos abiertos y gratuitos del programa COPERNICUS. Dentro de sus proyectos más destacados encontramos el Lobelia Air. Este comprende un sistema de visualización y reporte de concentraciones de materias particuladas (PM_{2,5} y PM₁₀) y de dióxido de nitrógeno, en tres ciudades; Madrid, Barcelona y Ámsterdam. En colaboración con la AGENCIA DE METEOROLOGÍA DE HOLANDA (KNMI), Lobelia Air se trata de un producto gratuito, que incluso ha venido siendo adoptado por entidades institucionales como el AYUNTAMIENTO DE MADRID en la búsqueda de solucionar problemas como el de la resolución detallada de la calidad de aire que con el sistema de estaciones de monitorización no pueden superar. Así como BREEZOMETER, Lobelia Air utiliza información de patrones de tráfico para determinar puntos calientes en estos

centros urbanos y proveer visualizaciones del tipo de mapas de calor para una asimilación ágil por los ciudadanos de las mismas.

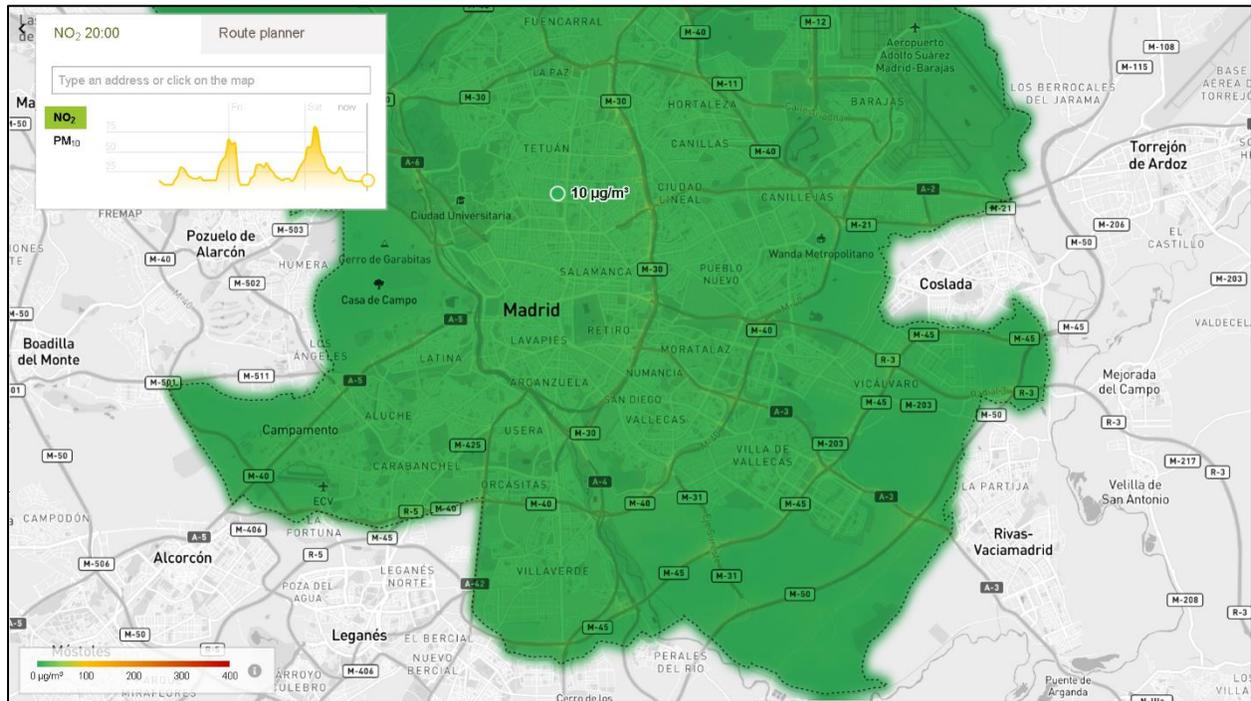


Fig. 11 Mapa de calidad del aire del sistema Lobelia Air sobre Madrid (Lobelia, Calidad del Aire en Madrid).

2. REVISIÓN BIBLIOGRÁFICA

Como hemos podido adelantar en múltiples ocasiones a lo largo de este trabajo, existe un inmenso valor agregado en conocer las concentraciones de los distintos contaminantes para todas las ubicaciones posibles en donde se puedan encontrar expuestos cualquier individuo o ser vivo, así como la posible vegetación, la cual también puede sufrir graves perjuicios como se ha indicado previamente. Así mismo, sabemos que excluyendo las posibles iniciativas privadas que se han mencionado, las cuales además prestan sus servicios a partir de un precio, en el caso de BREEZOMETER, las iniciativas gubernamentales se limitan a la recopilación de información en ubicaciones determinadas, dificultando el análisis de cualquier otra zona no cubierta por dicha red de vigilancia, por tanto, limitando la maximización del bienestar de toda su población.

Dentro de las posibles soluciones a esta problemática, encontramos la disciplina científica de la geostatística. Se trata de una rama de la estadística tradicional enfocada en conjuntos de datos

espaciales o espaciotemporales. De forma más precisa, esta disciplina “*comprende a un conjunto de herramientas y técnicas que sirven para analizar y predecir los valores de una variable que se muestra distribuida en el espacio o en el tiempo de una forma continua*” (MORAL GARCÍA, 2004, p. 79). Ahora bien, la interpolación espacial de datos forma parte de la geoestadística y “*se basa en el cálculo de los valores desconocidos de una variable espacial a partir de otros valores cuyo valor es conocido*” (QUESADA MUELAS, 2019, párrafo 1). Así mismo, dentro de esta extensa rama de la interpolación espacial de datos existen diversos métodos para el cómputo de esos valores desconocidos a partir de unos puntos muestrales, como lo puede ser el caso que aquí nos atañe; conocer las concentraciones de distintos contaminantes en áreas no cubiertas por las estaciones de monitorización, a partir de los datos recopilados en éstas.

Existen dos grandes clasificaciones de los métodos de interpolación espacial, las cuales se basan en las características de su naturaleza. Por una parte, encontramos los métodos de interpolación determinísticos, los cuales generan superficies continuas mediante el grado de similitud que, entre las observaciones registradas y aquellos puntos desconocidos, a través de su distancia. Dentro de esta clasificación encontramos los métodos de la ponderación de la distancia inversa o IDW, por sus siglas en inglés de *Inverse Distance Weighting*, y métodos como el *Spline*. Por otra parte, los métodos probabilísticos generan, igualmente superficies continuas de valores, pero en este caso aquellos valores desconocidos se consideran variables aleatorias y los datos disponibles se consideran “*realizaciones puntuales de un proceso aleatorio continuo*” (REBOLLO, Geoestadística, p. 2). En este tipo de métodos se lleva a cabo un análisis estructural de los datos para tomar en cuenta las posibles relaciones entre ellos de forma que no se puede garantizar la independencia entre los valores registrados ni en las predicciones a realizar sobre las variables aleatorias. Dentro de los métodos probabilísticos más comunes encontramos métodos como el Kriging en sus distintas complejidades: ordinario, universal, cokriging, entre otros. Como tendremos ocasión de explicar en secciones posteriores, estos métodos se diferencian en el grado de importancia que les atribuyen cada observación registrada para el cómputo de la nueva variable desconocida. Los métodos determinísticos, como el de la ponderación de la distancia inversa, simplemente toma en cuenta las distancias entre los puntos muestrales y los puntos a conocer para determinar el peso que debe tener cada valor registrado sobre el cómputo del nuevo punto, tomando menos en cuenta aquellos puntos muestrales más alejados de los puntos desconocidos. Por otra parte, los métodos probabilísticos como el Kriging toman en cuenta la variación espacial de los

puntos y sus posibles relaciones para determinar que peso atribuir a cada observación muestral en el cómputo del valor para un punto desconocido.

La bibliografía que detalla el uso de estos métodos de interpolación espacial es sumamente extensa. Encontramos estudios que utilizan estos métodos para el registro de un sinnúmero de materiales y componentes orgánicos en el subsuelo, de forma relacionada con el origen de métodos como el Kriging en la predicción de concentraciones de minerales en ciertos países africanos. Así mismo, encontramos estudios llevados a cabo en distintos países que tomen en cuenta la distribución de concentraciones de contaminantes y su relación con otros patrones espaciales como las características demográficas o económicas de la población.

De manera más relevante para nuestro trabajo, cabe mencionar el artículo de investigación elaborado por BHANDARI *et al.* en el 2022, el cual detalla los efectos de desastres naturales y, específicamente el del huracán Florencia, sobre la redistribución y aumento en la concentración de compuestos como el benceno, tolueno, etilbenceno y xileno (BTEX), así como de los contaminantes principales (materias particuladas, óxidos de nitrógeno, ozono troposférico, óxidos de azufre y monóxido de carbono), en el área estatal de CAROLINA DEL NORTE en los ESTADOS UNIDOS DE AMÉRICA. En este, los autores inician su investigación con la labor de recopilación de datos a lo largo de puntos muestrales en principales áreas del estado. Para ello, utilizan sensores especializados en el registro de compuestos BTEX por un período de 9 meses posteriores a la inundación e impacto del huracán Florencia (septiembre del 2018 y enero y mayo del 2019). Además, utilizan la plataforma de BREEZOMETER para obtener datos sobre los contaminantes principales y sobre la calidad del aire sobre los mismos puntos de muestreo de los compuestos BTEX, en el mismo espacio temporal, así como en el resto del territorio dadas las capacidades de resolución de la plataforma como hemos mencionado previamente.

Así mismos, los autores utilizan el método de interpolación espacial Kriging para “*modelar la distribución espacial de todos los compuestos químicos*” (BHANDARI *et al.*, 2022, p. 5). Con ello, y como procederemos posteriormente en este trabajo nosotros, los autores parten de los datos de BTEX obtenidos a partir de sus registros muestrales y utilizan dicho método de interpolación espacial para predecir el valor de BTEX a una resolución de 500 metros a lo largo de todo el estado, durante el espacio temporal mencionado. Aún más, en este trabajo los autores contrastan los valores obtenidos mediante el método Kriging con las ubicaciones de estaciones de servicio

afectadas por las inundaciones e impactos del huracán, observado gran correlación en el incremento de las concentraciones de estos compuestos químicos BTEX y su proximidad con las estaciones de servicios, probablemente fuentes de emisión de estos. Por último, los autores arriban a conclusiones interesantes y específicas en el señalamiento de un aumento significativo del compuesto benceno y de las materias particuladas PM_{2,5} inmediatamente después de las inundaciones ocasionadas por el huracán, sobrepasando los límites regulatorios de las agencias estatales sobre la calidad del aire, así como un lento y prolongado descenso y disipación de estos contaminantes en meses posteriores (BHANDARI *et al.*, 2022).

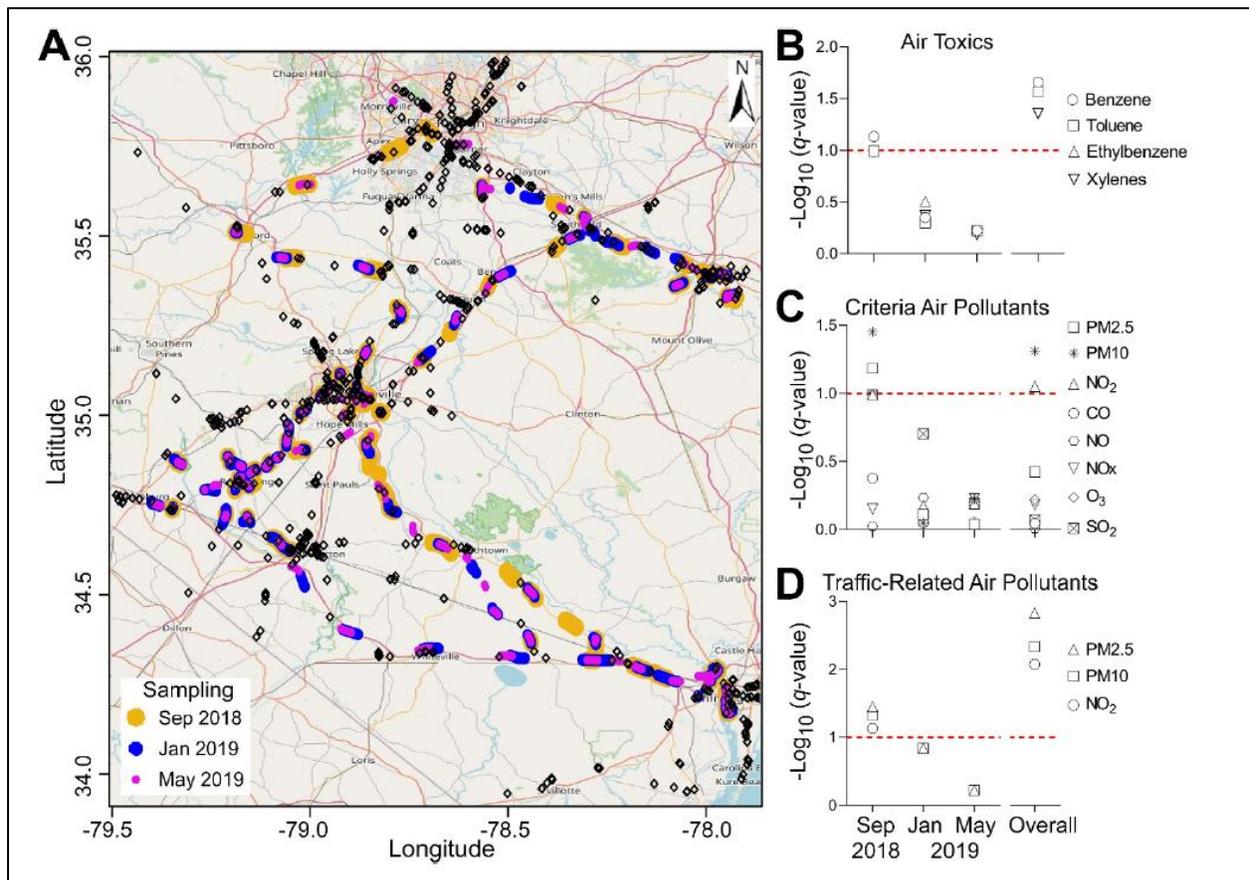


Fig. 12 Asociación especial de las concentraciones de compuestos químicos BTEX y de contaminantes principales con estaciones de servicio (Bhandari *et al.*, 2022, p. 12).

Además de este artículo de investigación, resulta de vital importancia exponer en este apartado el trabajo elaborado por CAÑADA TORRECILLA *et al.* titulado INTERPOLACIÓN ESPACIAL Y VISUALIZACIÓN CARTOGRÁFICA PARA EL ANÁLISIS DE LA JUSTICIA AMBIENTAL: ENSAYO

METODOLÓGICO SOBRE LA CONTAMINACIÓN POR PARTÍCULAS ATMOSFÉRICAS EN MADRID. En este, investigadores de la UNIVERSIDAD AUTÓNOMA DE MADRID exponen la forma en la que ciertos contaminantes principales se encuentran distribuidos en el AYUNTAMIENTO DE MADRID y como dicha distribución puede causar un perjuicio mayor para poblaciones con menores recursos, expuestas en mayor medida a peores calidades del aire en zonas urbanas determinadas de la provincia (CAÑADA TORRECILLA *et al.*, 2010).

Para ello, los autores han utilizado datos de contaminantes principales a través de las RED DE VIGILANCIA DE LA CALIDAD ATMOSFÉRICA del AYUNTAMIENTO DE MADRID, integrada, como ya hemos visto previamente, por más de 20 estaciones de monitorización distribuidas a lo largo del territorio. De manera más precisa, éstos se enfocaron en “*el número de días con concentraciones de PM10 superiores a 50 µg/m3*” el cual se corresponde con el valor límite diario que impone la legislación española, como hemos visto previamente, sobre este particular contaminante. Así mismo, han utilizado un período de evaluación comprendido entre 1997 y 2006 para su estudio.

De forma similar a como se expondrá posteriormente, los autores de este trabajo han utilizados métodos de interpolación espacial para obtener datos de concentración de PM10 en aquellas zonas del AYUNTAMIENTO DE MADRID que no se encuentran cubiertas por las estaciones de monitorización presentes. Aún más, los autores contrastan los resultados obtenidos a partir de la utilización de la ponderación de la distancia inversa, como método determinístico, y el kriging ordinario como método probabilístico, en cuanto a su precisión de predicción y demás efectos sobre la recopilación de datos. Así mismo, los autores utilizan la plataforma privada de ArcGIS para la elaboración de estos modelos, así como para las visualizaciones que exponen en el trabajo (CAÑADA TORRECILLA *et al.*, 2010).

Por otra parte, una vez obtenidas las frecuencias de días en los que se rebasa el umbral máximo de concentraciones de PM10, los autores utilizan los valores de renta per cápita en las distintas secciones censales como aproximación al nivel socioeconómico de las poblaciones respectivas, y para poder posteriormente contrastar la calidad del aire en los puntos más y menos favorecidos económicamente.

Dentro de sus conclusiones, el grupo de trabajo expone la ligera mejora que experimenta el método kriging en contraposición del de ponderación de la distancia inversa en cuanto al error cuadrático medio sobre los puntos desconocidos. Por otra parte, y atribuyendo a esta mejora sobre el error cuadrático medio, señalan que el método kriging suaviza mucho más extensamente los valores predictivos de forma que el mínimo valor predicho resulta mayor que el mínimo predicho por la ponderación de distancia inversa. De forma similar, el máximo predicho por kriging será menor en magnitud que el máximo computado por el IDW. Aún más, a pesar de obtener un mejor error cuadrático medio en el kriging, para muchas de las observaciones probadas el IDW muestra ser mejor método de interpolación espacial, precisamente por este suavizado que realiza el kriging. Finalmente, se concluye que existe una ligera variación en la consideración de calidad del aire para las distintas secciones censales agrupadas por renta per cápita dependiendo del método de interpolación espacial que se utilice, siendo el IDW más local y severo en promedio. Así mismo, no se puede concluir que haya una consistencia en la calidad del aire por zonas de nivel socioeconómico en el Ayuntamiento de Madrid, teniendo tanto las rentas más bajas como las más altas el mayor número de días superando el umbral de contaminación máximo de PM10 (CAÑADA TORRECILLA *et al.*, 2010).

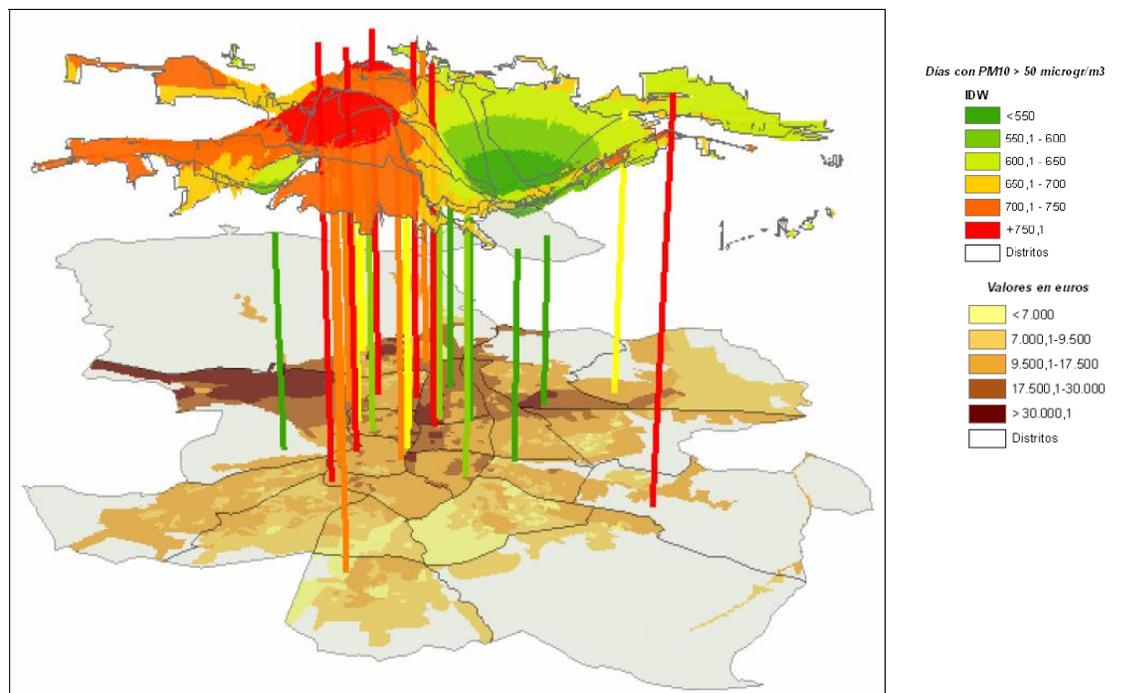


Fig. 13 Representación de las zonas de renta (abajo) y los días estimados (con IDW-PID) de excesivas PM10 (arriba) (CAÑADA TORRECILLA *et al.*, 2010, p. 708)

CAPÍTULO III. ANÁLISIS DE LA RED DE VIGILANCIA DEL AYUNTAMIENTO DE MADRID.

1. ESTRUCTURA DE LOS DATOS. ANÁLISIS EXPLORATORIO.

Como hemos adelantado previamente, nuestro análisis partirá de los datos recogido por la RED DE CALIDAD DEL AIRE DE LA COMUNIDAD DE MADRID y del AYUNTAMIENTO DE MADRID, ambas integradas al sistema del ÍNDICE NACIONAL DE CALIDAD DEL AIRE (ICA) y compuesto por 48 estaciones de monitorización distribuidas en toda la provincia de Madrid. Para el acceso a los mismos, encontramos que el portal de la red de vigilancia de la COMUNIDAD DE MADRID nos redirige al portal de descarga masiva de datos abiertos de la comunidad, disponible en el siguiente enlace:

https://datos.comunidad.madrid/catalogos?q=aire&sort=score+desc%2C+metadata_modified+desc

Su descarga viene un formato csv (*comma separated value*) para cada una de las estaciones que forma parte de la red de vigilancia. La estructura de los mismos trata de datos en serie temporal incluyendo tanto año, mes y día como las 24 horas del día para las cuales se registran las concentraciones de los determinados contaminantes por cada estación de monitorización; propiamente de: Benceno, Hidrocarburos no metánico, Monóxido de carbono, Ozono troposférico, Tolueno, Dióxido de azufre, Hidrocarburos totales, Monóxido de nitrógeno, Partículas en suspensión (PM10), Partículas en suspensión (PM2,5), Dióxido de nitrógeno, Metaparaxileno, y Óxido de nitrógeno.

Estas series temporales, detalladas hasta un nivel horario, incluyen, además de las concentraciones de dichos contaminantes por cada hora, siete (7) otros parámetros de interés. Estos son:

- MAGNITUD: Valores de tipo numérico discreto que representan el contaminante en particular que se mide en cada fila (año, mes y día) para toda hora del día en cada estación.
- DESCRIPCIÓNMAGNITUD: Valores de tipo categórico (nominales) que toman los nombres de los contaminantes propiamente (Dióxido de nitrógeno, e.g.), para cada codificación de magnitud.
- CÓDIGO NACIONAL: Valores de tipo numérico discreto que codifican la estación de monitorización en número de 8 dígitos.

- NOMBRE ESTACIÓN: Valores de tipo categórico (nominales) que establecen el nombre atribuido a la estación de monitorización sobre la cual se recogen los datos.
- LATITUD: Valores de tipo numérico continuo que determinan la latitud, coordenadas Norte-Sur, en la que se ubica la estación de monitorización sobre la cual se recogen los datos.
- LONGITUD: Valores de tipo numérico continuo que determinan la latitud, coordenadas Norte-Sur, en la que se ubica la estación de monitorización sobre la cual se recogen los datos.
- MEDIA DIARIA: Valores de tipo numérico continuo que representa la media aritmética de los valores recogidos por la estación de monitorización para cada contaminante tomando en cuenta las 24 horas del día registradas.

	A	B	C	D	E	F	G	H	I	J	K	L	M
	MAGNITUD	DESCRIPCIONMAGNITUD	CODIGO NACIONAL	NOMBRE ESTACION	LATITUD	LONGITUD	ANO	MES	DIA	H01	H02	H03	H04
1	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	1	25	20	18	5
2	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	2	1	1	1	1
3	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	3	1	3	3	4
4	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	4	1	1	1	1
5	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	5	54	58	46	45
6	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	6	35	92	29	28
7	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	7	1	1	1	1
8	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	8	1	1	1	1
9	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	9	2	3	1	1
10	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	10	2	2	1	1
11	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	11	1	1	1	1
12	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	12	1	1	1	1
13	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	13	1	8	3	2
14	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	14	81	59	22	40
15	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	15	25	18	8	9
16	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	16	1	1	1	1
17	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	17	1	1	1	1
18	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	18	1	1	1	1
19	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	19	22	9	3	3
20	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	20	7	6	4	4
21	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	21	1	1	1	1
22	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	22	1	2	4	2
23	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	23	1	1	1	1
24	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	24	1	1	1	1
25	7	Monóxido de nitrógeno	28009001	ALGETE	-3.5032861	40.5998194	2012	1	24	1	1	1	1

Fig. 14 Datos de la estación de monitorización de Algete.

Luego de la extracción de los datos de cada una de las estaciones de monitorización de la COMUNIDAD DE MADRID y del AYUNTAMIENTO DE MADRID, hemos procedido a su carga en el entorno de programación de R STUDIO, en el cual implementaremos el resto del análisis e investigación para las finalidades de este Trabajo de Fin de Grado, comenzando, necesariamente, por un análisis exploratorio de los mismos.

En primer lugar, resulta imperativo conocer la distribución estadística de los distintos contaminantes, tomando en cuenta todos los valores registrados en las distintas estaciones de monitorización, en términos de sus respectivos valores medios, ubicación de cuartiles, mínimos y

máximos, detalles sobre la varianza de los mismos, así como la presencia de valores atípicos (*outliers*). Para ello, empezamos por elaborar un gráfico comparativo de diagramas de cajas y bigotes (boxplots) para cada uno de los contaminantes considerando los valores de sus medias diarias, visto en la figura 15.

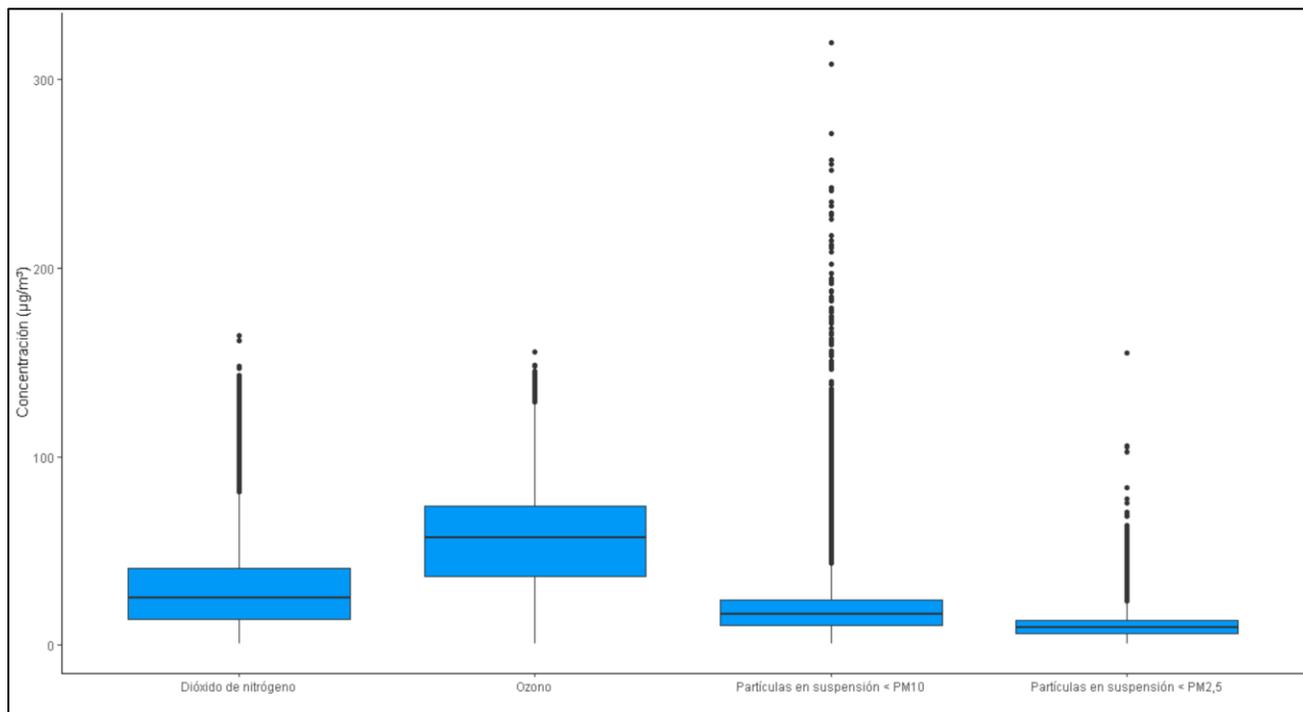


Fig. 15 Gráfico de cajas y bigotes de los contaminantes principales.

Como podemos observar, cada uno de los contaminantes registrados contiene un ingente número de valores atípicos, teniendo las materias particuladas PM10 el mayor número de ellos. Así mismo, vemos que, de forma relativa, los valore extremos de este contaminante también muestran ser los más desproporcionados en comparación con los valores habituales registrados. Luego de proceder a la eliminación de estos valores extremos elaboramos una comparativa de estadísticas básicas sobre cada contaminante visto en la figura 16.

Medida	Dióxido de nitrógeno ($\mu\text{g}/\text{m}^3$)	Ozono ($\mu\text{g}/\text{m}^3$)	Partículas en suspensión (PM10) ($\mu\text{g}/\text{m}^3$)	Partículas en suspensión (PM2,5) ($\mu\text{g}/\text{m}^3$)
Media	28.025	55.177	17.402	9.715
Mediana	24.583	57.333	15.833	8.917
Máximo	81.167	128.95	43.667	23.333
Mínimo	1.000	0.994	1.000	1.000

Cuartíl 1 (0.25)	13.583	36.748	10.542	6.125
Cuartíl 3 (0.75)	39.667	73.667	22.812	12.542
Desviación Típica	18.115	24.989	8.790	4.624
Coefficiente de Variación (%)	64.639	45.288	50.515	47.599

Fig. 16 Tabla descriptiva de valores estadísticos de cada contaminante principal

Con estos datos podemos ver claramente como los niveles de concentración medidos en las mismas unidades, microgramos por metro cúbico ($\mu\text{g}/\text{m}^3$), resultan incomparables entre si dados los rangos habituales de cada uno. De esta forma, podemos ver como las medias y medianas de cada contaminante se encuentran valores muy alejados uno del otro. Para ello, procedimos a la introducción de medidas como el Coeficiente de Variación de Pearson, el cual determina la relación entre la desviación típica de cada variable y su media, indicando, por tanto, la variación relativa de los datos, en términos porcentuales, y siendo posible su comparación entre variables. Podemos ver aquí como el dióxido de nitrógeno muestra tener la mayor variabilidad en cuanto a sus concentraciones medias diarias registradas con un 64,7% aproximadamente.

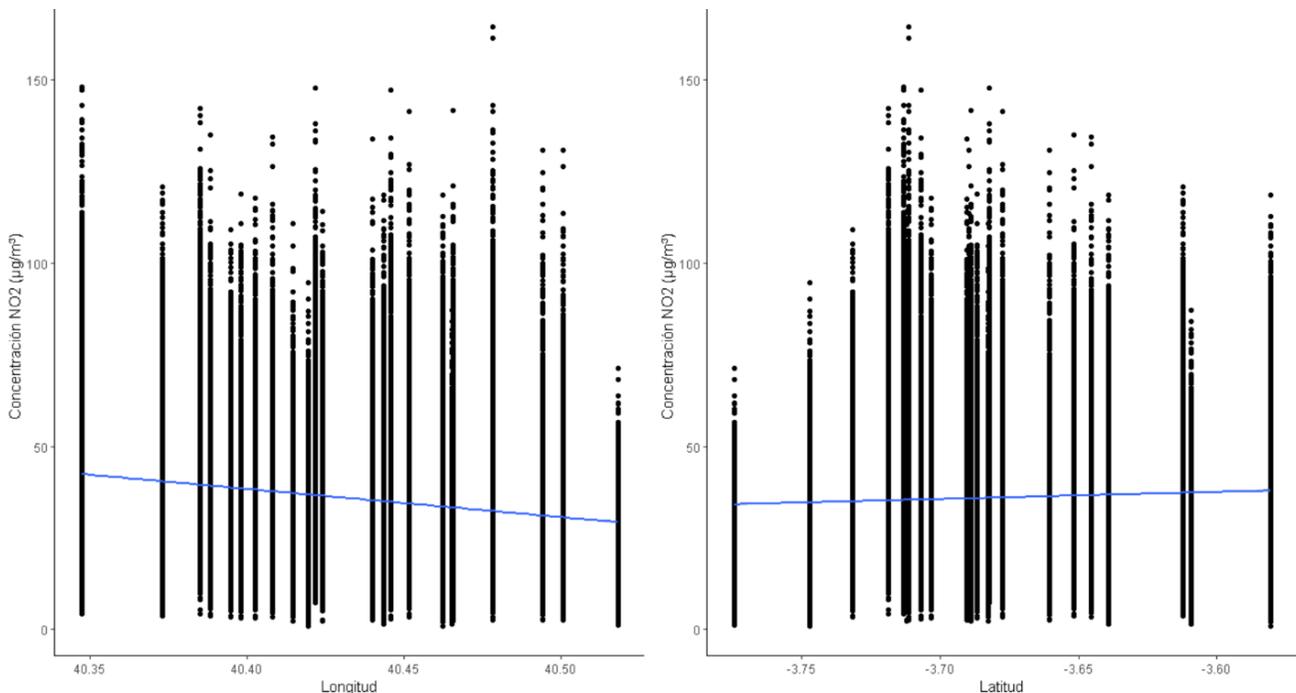
De manera similar, podemos observar algunas discrepancias entre las medias y medianas observadas para cada contaminante. Las diferencias entre estos datos para el NO₂, PM₁₀ y PM_{2,5} sugieren una ligera asimetría de los datos hacia a la derecha, ya que su media se encuentra por encima de la mediana. Por el contrario, la media y mediana del O₃ sugiere una menor asimetría de los datos, pero en este caso hacia la izquierda, siendo la media menor que la mediana. Al medir sus Coeficientes de asimetría de Fisher, pudimos observar una efectiva ligera asimetría a la derecha de aproximadamente 0,7 para los primeros tres contaminantes mencionados. En cambio, pudimos entender una aceptable simetría en la distribución de los datos del O₃ teniendo un coeficiente de asimetría igual a -0,12.

Por otra parte, resulta importante mencionar la existencia de un gran número de valores perdidos (NAs) en los datos de cada estación de monitorización. Se da el caso de tener hasta 740 valores perdidos en alguno de los datos de las estaciones de monitorización. Sin embargo, y, a pesar de este ser el máximo número de NAs para una determinada estación, la media de valores perdidos en cada estación es de 116.4. Esto no conllevaría mayor problemática debido al elevado número de observaciones (filas) que posee cada una de las estaciones en sus datos; 21.325,36 observaciones en promedio. Por ello, los valores perdidos representarían únicamente un 0,55% de los datos de cada estación aproximadamente. Como hemos anticipado en capítulos anteriores, esta

gran cantidad de valores perdidos se puede atribuir a algún fallo puntual de los sensores de cada estación de monitorización, así como a los posibles niveles imperceptibles de determinados contaminantes en días concretos.

Al tratar con datos georreferenciados, nos resulta posible visualizar la posible variabilidad de los datos a lo largo del espacio en el cual son registrados. No solo se trata de información interesante para entender las tendencias del valor promedio de los distintos contaminantes, sino que, como se verá más adelante, resulta imperativo para determinar los mejores métodos de interpolación espacial a aplicar en dicha configuración del territorio. Para ello, y centrándonos en nuestro contaminante de interés (NO₂), en la siguiente figura (17) se podrá observar, por una parte, la dispersión de los valores medios diarios de concentración de NO₂ en la dirección este-oeste, y, por otra parte, esta misma dispersión en la dirección norte-sur. En cuanto a la dispersión sobre la latitud, se puede ver cómo, a pesar de la precisa longitud o latitud en la cual se registren los datos, el valor promedio de este contaminante no muestra mayor tendencia, indicando una aparente estacionariedad de la concentración de NO₂ a lo largo de la región de estudio. Por otra parte, vemos como la concentración de NO₂ sobre la longitud va disminuyendo a medida que nos dirigimos en dirección este (mayor longitud y mayores áreas verdes en Madrid), experimentando una ligera tendencia y determinando la necesidad de tomar en cuenta la configuración espacial

Fig. 17. Gráfico de dispersión de NO₂ a lo largo de la región en cuanto a longitud (izquierda) y latitud(derecha).



para las interpolaciones espaciales a calcular, abogando entonces, a priori por un método como el kriging.

Seguidamente, sería interesante entender algunas estadísticas simples para los valores registrados en cada estación dependiendo de la hora en la cual fueron registrados. Con un análisis de la distribución de todos los valores de cada estación de monitorización, podemos ver a través de los cuartiles de los mismos, que, en media (percentil 50), las horas con concentraciones más elevadas corresponden a aquellas horas de 09h, 10h, 20h, 21h y 22h. Por otra parte, al ver los valores máximos para cada hora, confirmamos que las horas nocturnas (20h, 21h y 22h) son, en media, mucho más elevados que los registrados a cualquier otra hora (aproximadamente 233% del valor medio máximo de todas las horas), tomando en cuenta todo tipo de contaminante. Probablemente la explicación de este fenómeno encuentre resguardo en el elevado número de tráfico de coches que concurren cerca de las áreas de cobertura de dichas estaciones de monitorización. Estos resultados pueden verse en la figura 18 a continuación.

Hora	Media ($\mu\text{g}/\text{m}^3$)	Cuart. 1 (0.25) ($\mu\text{g}/\text{m}^3$)	Cuart. 3 (0.75) ($\mu\text{g}/\text{m}^3$)	Máximo ($\mu\text{g}/\text{m}^3$)
H01	23.6472	2.77234	31.44335	533.1277
H02	21.23306	2.52617	28.66915	482.0638
H03	19.29149	2.397872	25.97181	424.766
H04	18.03355	2.363404	24.24128	395.766
H05	17.73022	2.579787	24.36787	375.6383
H06	19.09561	3.091277	26.46809	365.5319
H07	22.47107	3.86383	30.73596	407.0426
H08	27.42088	4.876596	36.75809	531.8723
H09	30.34704	5.373989	39.98904	680.2979
H10	29.20355	5.453617	38.30252	616.6809
H11	26.50313	5.004255	35.37229	509.617
H12	25.01703	4.630443	33.42027	448.5785
H13	24.29107	4.170922	32.39731	388.0426
H14	23.79651	3.763032	31.275	386.2553
H15	23.36092	3.429787	30.36771	375.6436
H16	23.03624	3.304255	29.82559	368.4468
H17	23.48281	3.31844	31.01362	395.3404
H18	25.52759	3.598032	34.20156	562.7447
H19	28.81472	3.876596	38.14691	744.617
H20	31.57513	4.276596	40.8509	858.7021
H21	32.77703	4.37266	41.96255	875.6809
H22	32.1178	4.126383	41.33309	820.5106
H23	29.86368	3.715213	38.47638	712.6809
H24	26.65535	3.203511	34.92957	619.7234

MEDIA DIARIA	25.22053	5.674825	34.75853	364.1215
--------------	----------	----------	----------	----------

Fig. 18 Análisis de la distribución de valores sobre la concentración total de contaminantes por hora.

Así mismo, es importante observar exactamente cuáles contaminantes se ven registrados en las distintas estaciones de monitorización que tenemos a nuestra disposición. La propia CONSEJERÍA DE MEDIO AMBIENTE, VIVIENDA Y AGRICULTURA, encargada de gestionar la recopilación de datos acerca de estos contaminantes en toda la COMUNIDAD DE MADRID, expone que no todas las estaciones de monitorización recopilan, de forma automática, todos los contaminantes anteriormente expuestos. En cambio, se ha distribuido la recopilación de datos sobre contaminantes como SO₂, PM₁₀, PM_{2,5} y los BTX en las zonas que a su parecer han tenido la mayor incidencia e importancia en estudiar. Por tanto, de nuestros propios datos hemos podido observar que los únicos contaminantes que encontramos presentes en todas las estaciones de monitorización son el dióxido de nitrógeno (NO₂) y el óxido de nitrógeno (NO). Por su parte, el ozono troposférico (O₃) se encuentra registrado en 37 de las 47 estaciones recopiladas. Así mismo, las materias particuladas PM₁₀ se encuentran registradas en 31 de las estaciones disponibles y las inferiores a 2,5 micrómetros de diámetro (PM_{2,5}) se encuentran presentes en 20 de estas 47.

Finalmente, es importante mencionar que, debido, nuevamente, a la cantidad de valores perdidos y demás problemáticas características de las estaciones de monitorización de la red de vigilancia en cuestión, no hay una uniformidad temporal sobre todos los datos extraídos de todas las estaciones de monitorización. El espacio temporal que abarca cada conjunto de datos comprende los años desde el 2012 hasta el 2022, sin embargo, solo 24 estaciones de las 47 totales contienen datos del periodo 2012-2016. No obstante, las observaciones desde el 2017-2021 se muestran suficientes para nuestro análisis, encontrándose incluido este período para todas las estaciones de monitorización utilizadas.

2. VISUALIZACIONES GEORREFERENCIADAS EXPLORATORIAS

Las visualizaciones elaboradas en este Trabajo de Fin de Grado se han realizado, principalmente, mediante las librerías *mapview*, *terra*, *ggplot2*, *rgeos* y *rgdal*. Cada una de estas presenta capacidades de visualización muy potentes y, con la excepción de *ggplot2*, poseen ciertas especialidades para el tratamiento de datos georreferenciados como es el caso que nos atañe.

En primer lugar, resultaba necesario acceder a los datos georreferenciados de toda la COMUNIDAD DE MADRID. Usualmente, este tipo de datos sobre el cual se podrían elaborar visualizaciones a partir de datos de latitud y longitud se encuentran almacenados en formato shapefile (.shp). Para ello, procedimos al portal de la COMUNIDAD DE MADRID y su apartado de cartografía y temática, el cual contiene un sinfín de recursos digitales de toda la provincia de Madrid, y entre ellos los mapas digitales georreferenciados de tipo shapefile necesarios para nuestras visualizaciones, disponible en el siguiente enlace: <https://www.comunidad.madrid/servicios/urbanismo-medio-ambiente/cartografia-topografica-tematica>.

Una vez descargado estos archivos, se procedió al análisis y limpieza de estos en el entorno de R STUDIO. De manera más precisa, la mayor complejidad en la adecuación de los datos obtenidos consistió en la determinación del sistema de referencias de coordenadas (CRS, en inglés, por sus siglas, *Coordinate Reference System*). Como bien se sabe, la manera tradicional de representar la tierra se corresponde con un globo. No obstante, la proyección de la tierra en dos dimensiones conlleva grandes problemas si se mantiene esta forma tradicional de globo para representarla. Para solventar esta problemática, se han venido desarrollando distintas formas de proyectar los datos de la tierra sobre dos dimensiones, aunque estas soluciones divergen en ciertos aspectos y resulta vital entender el sistema de proyecciones correcto sobre el cual se proyecta el mapa trabajado para obtener visualizaciones correctas y precisas (QGIS, Proyección cartográfica en detalle, párrafo 2). Por otra parte, los sistemas de referencias de coordenadas (CRS) definen “*cómo el mapa proyectado bidimensional en su SIG se relaciona con lugares reales en la tierra*” (QGIS, Proyección cartográfica en detalle, párrafo 2). De forma más precisa, el CRS utiliza una serie de valores codificados que permiten, para cada proyección de la tierra, ubicar puntos precisos en el mapa. El sistema de referencias de coordenadas más popular obedece al que utiliza la latitud y longitud siguiendo el sistema WSG 84, EPSG:4326 (*World Geodetic System 1984*).

En nuestro caso, la cartografía georreferenciada que pudimos obtener del portal de la COMUNIDAD DE MADRID sigue el sistema ETRS89 / UTM zone 30, EPSG:25830, denominado *European Terrestrial Reference System 1989*. Este sistema de proyección de mapas permite un nivel de resolución sobre áreas de superficie relativamente pequeñas, las cuales bajo cualquier otro sistema de proyección sería imposible de conseguir para países en Europa, entre ellos, España. Ahora bien, los datos obtenidos de la red de vigilancia de la calidad del aire incluyen las latitudes y longitudes

de cada estación en concordancia con el sistema más popular de WSG 84, EPSG:4326, y por ello resultó necesario transformar dichas coordenadas al sistema de referencias de coordenadas (CRS) correspondiente al del mapa cartográfico que utilizaremos; ETRS89 / UTM zone 30, EPSG:25830.

Con todo ello, resulta posible elaborar algunas visualizaciones exploratorias de los datos, como lo pueden ser un simple mapa de la COMUNIDAD DE MADRID indicando en el mismo la ubicación de cada una las estaciones de monitorización utilizadas y un gradiente de color correspondiente con los valores de media diaria histórico tomando en cuenta todos los contaminantes registrados, visto en la figura 19 a continuación. Como se puede ver en esta visualización, una gran parte de las estaciones de monitorización se encuentran muy cerca una de la otra. Esto se debe a la red de vigilancia de la calidad del aire del AYUNTAMIENTO DE MADRID, la cual se encuentra gestionada por este mismo ente local y que por su densidad poblacional, resulta necesario obtener la mayor precisión sobre áreas más reducidas.

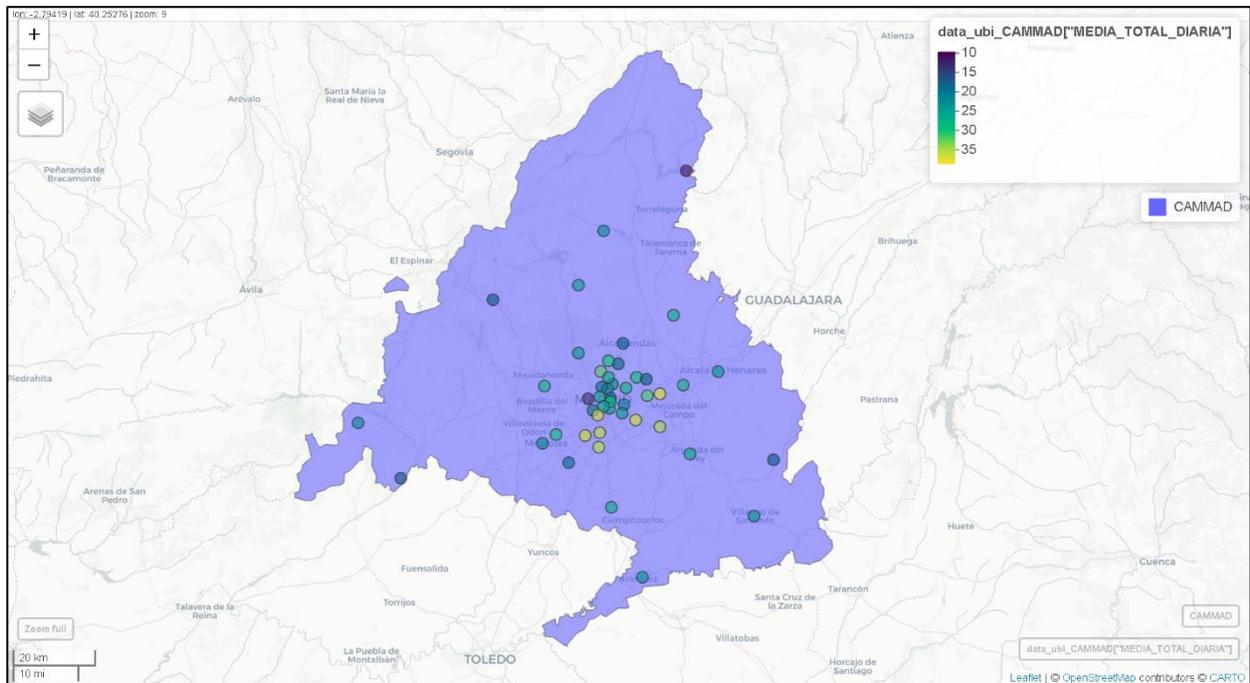


Fig. 19 Estaciones de Monitorización de la Comunidad de Madrid y sus niveles de media diaria históricos.

Así mismo, puede resultar interesante realizar una visualización sobre una observación que ya hemos comentado previamente en cuanto a las diferencias horarias y sus concentraciones de contaminantes. Para ello, procederemos a plasmar los valores máximos registrados en cada

estación de monitorización sobre el total de contaminantes para ver si existe alguna relación espacial en este aspecto.

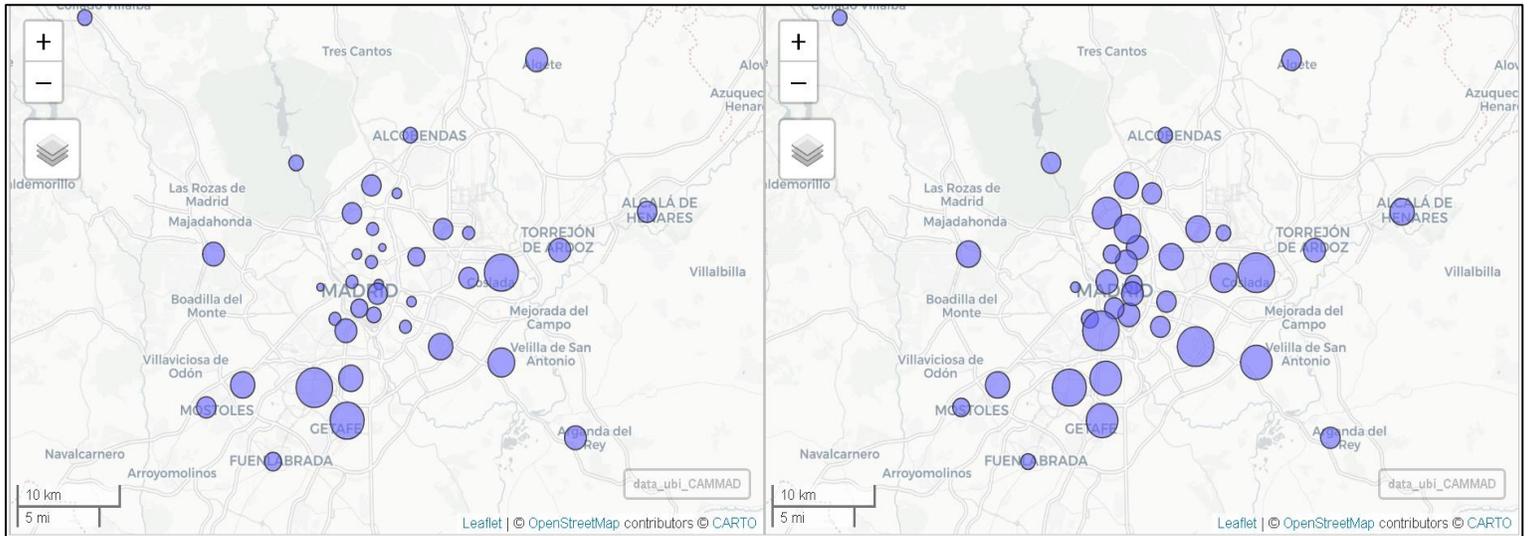


Fig. 20 Vista comparativa de concentración (tamaño) de todos los contaminantes entre las 06h (izquierda) y las 21h (derecha) en las estaciones de monitorización.

Como podemos observar en la figura 20, existe una gran variación entre los valores registrados en las estaciones de monitorización al comparar sus valores en las horas 06h y 21h. Por una parte, considerando únicamente los valores registrados a las 06h, observamos unos pocos puntos de gran tamaño (concentración) ubicados a los alrededores del área metropolitana de la ciudad, registrando el resto de estos unos valores muy bajos en comparación. Por otra parte, al observar los datos recopilados a las 21h, seguimos observando valores sumamente altos en las mismas estaciones de monitorización que el gráfico de la izquierda, pero también notamos un aumento sistemático de los puntos ubicados dentro del área metropolitana de Madrid.

Por último, y, como aproximación exploratoria, resulta interesante observar las posibles relaciones que existen entre las categorías en las cuales clasifican los entes gubernamentales las distintas estaciones de monitorización y sus respectivos valores medios diarios para las concentraciones registradas de los 4 contaminantes principales (NO₂, O₃ y PMs). Tanto la COMUNIDAD DE MADRID como el AYUNTAMIENTO DE MADRID, utilizan 6 categorías que aportan más información sobre las localidades en donde se encuentran ubicadas estas estaciones de monitorización. De manera más específica, éstas se abarcan en dos grupos.

En primer lugar, encontramos el tipo de área propiamente en donde se encuentra la estación de monitorización, obedeciendo a características de desarrollo urbanístico y densidad poblacional. En este grupo encontramos tres clasificaciones; áreas rurales, áreas urbanas y áreas suburbanas. Por otro parte, el segundo grupo, denominado tipo de estación, nos aporta la información relativa a las principales fuentes de contaminación que se encuentran más próximas a cada estación y, por ello, sus clasificaciones consisten en áreas industriales, áreas de tráfico y áreas de fondo, estas últimas referenciando la actividad humana diaria como como la principal fuente de contaminación.

A continuación, podremos observar dos visualizaciones que comprenden gráficos de barras agrupadas que recogen la información previamente descrita. En la figura 21 podemos observar la distribución de los valores medios para cada contaminante principal y su comparación entre las áreas de fondo, áreas industriales y áreas de tráfico. Rápidamente, podemos notar como para casi todos los contaminantes, con la excepción del ozono troposférico (O₃), las estaciones de monitorización ubicadas en áreas de tráfico se ven representadas más que cualquier otro tipo de estaciones. Además, de manera sorprendente, las estaciones de monitorización ubicadas en áreas industriales registran las concentraciones más reducidas para todos los contaminantes abarcados en este gráfico.

Así mismo en la figura 22, observamos un gráfico de barras agrupadas que contiene información similar. No obstante, en este la agrupación corresponde a las características de densidad poblacional y desarrollo urbanístico de las respectivas estaciones de monitorización. En esta, podemos ver como el ozono troposférico se encuentra en mayor proporción en áreas no urbanizadas (rurales y suburbanas). Por otra parte, los mayores niveles de concentración del dióxido de nitrógeno se encuentran registrados en estaciones de monitorización ubicadas zonas urbanas. Finalmente, las materias particuladas de ambas extensiones se ven similarmente distribuidas entre las distintas áreas de la COMUNIDAD DE MADRID, con una ligera preponderancia en zonas urbanas y suburbanas.

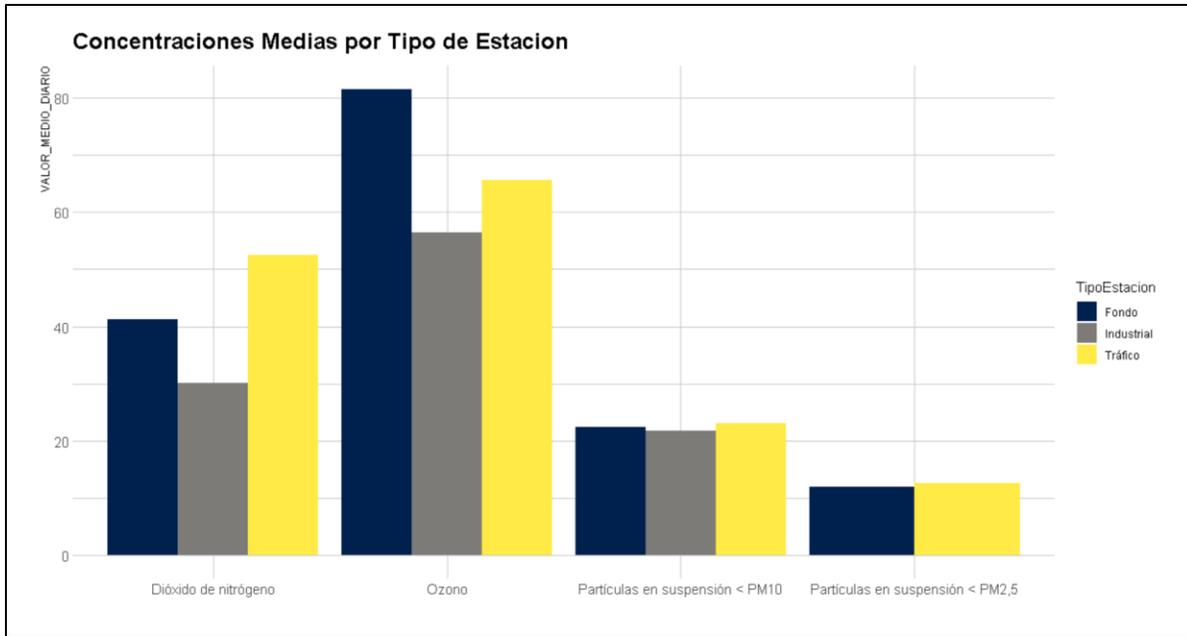


Fig. 21 Gráfico de barras agrupadas detallando las concentraciones medias de cuatro contaminantes principales en relación con las fuentes de contaminación de mayor proximidad.

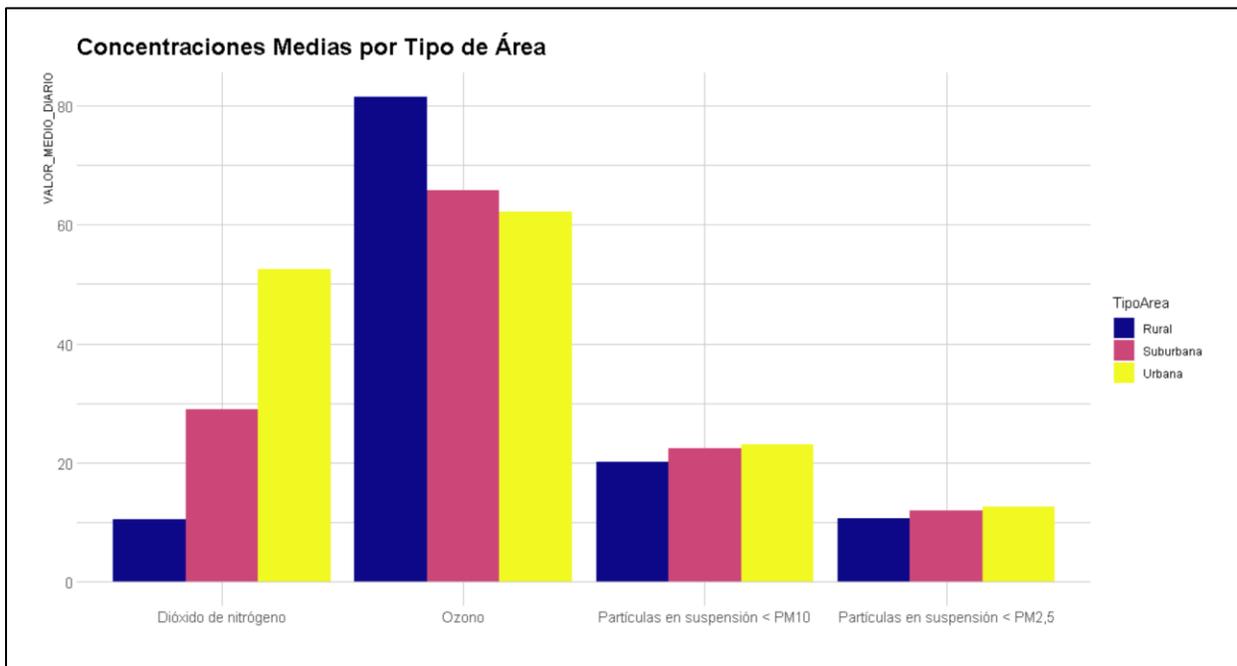


Fig. 22 Gráfico de barras agrupadas detallando las concentraciones medias de cuatro contaminantes principales en relación con el tipo de área en el cual se registran las mismas.

3. ANÁLISIS DE MODELOS DE INTERPOLACIÓN ESPACIAL

3.1. Ponderación de la Distancia Inversa (IDW)

Como tuvimos ocasión de adelantar en secciones anteriores de este Trabajo de Fin Grado, los métodos de interpolación espacial comprenden el conjunto de herramientas que ofrece la disciplina de la geoestadística para aproximarnos a valores desconocidos a partir de valores muestrales. Estos métodos se clasifican en métodos determinísticos o métodos probabilísticos dependiendo de la consideración que se le atribuyendo a los valores desconocidos como variables aleatorias o no.

Dentro de los métodos determinísticos, el más utilizado es el de la ponderación de la distancia inversa (IDW). Este método *“supone explícitamente que los elementos cercanos entre sí son más parecidos que los que están más alejados”* (ArcGIS Pro, Cómo funciona la interpolación de la distancia inversa, párrafo 1). Por ello, este método de interpolación espacial asume una influencia local de los valores muestrales sobre los valores predichos. Así mismo, esta influencia local se ve disminuida gradualmente, de tal manera que a mayor distancia menor será el grado de información que aporta el valor del punto muestral sobre el punto desconocido.

Esta influencia local de los puntos muestrales sobre los puntos desconocidos proviene del grado de ponderación que viene atribuido sobre el valor muestral en el contexto del cómputo del nuevo valor a conocer. Es decir, la importancia de cada dato muestral en el cómputo del nuevo dato, en nuestro caso, nueva concentración de un determinado contaminante se ve disminuida a medida que este punto muestral se aleja del punto desconocido. Por lo tanto, *“las ponderaciones son proporcionales a la inversa de la distancia (entre el punto de datos y la ubicación de predicción) elevada al valor de potencia ‘p’”* (ArcGIS Pro, La función Potencia, párrafo 1). Una vez computadas las potencias “p” para cada par punto muestral-punto desconocido, podemos efectuar la suma ponderada de los datos observados en los puntos muestrales, y así conocer el valor estimado para el punto desconocido.

$$Z(S_0) = \sum_{i=1}^N \omega_i * Z(S_i) \quad (1)$$

La representación matemática arriba descrita se ve evidenciada en la formula (1), en donde $Z(S_0)$ representa el valor desconocido a estimar sobre el punto S_0 . Seguidamente, el otro lado de la igual se compone de $\sum_{i=1}^N \omega_i * Z(S_i)$, un sumatorio que recorre N, el conjunto de puntos muestrales (S_i),

empezando desde el primero de ellos. Su recorrido suma todos los productos de las ponderaciones de cada punto muestral (ω_i) y el valor registrado en cada uno de ellos, $Z(S_i)$.

Ahora bien, y de manera más importante, las ponderaciones de cada punto muestral (ω_i), obedecen a la formula (2) expuesta debajo. En ella, se puede claramente como el cómputo de cada ponderación se obtiene a partir de una fracción. En su numerador encontramos la distancia que separa a la pareja punto muestral-punto desconocido (d_{i0}^{-p}), y elevada esta a la potencia “p” con signo negativo. En su denominador encontramos un sumatorio que recorre, nuevamente, todos los puntos muestrales empezando por el primero y asume todas las distancias entre los dichos puntos y el punto desconocido en cuestión, elevado a su vez a la potencia “p” con símbolo negativo.

$$\omega_i = \frac{d_{i0}^{-p}}{\sum_{i=1}^N d_{i0}^{-p}} \quad (2)$$

Ahora bien, resulta vital comprender que el valor asignado a la potencia “p” trata de un valor arbitrario cuya determinación tiene importantes consecuencias. Dado su símbolo negativo, este se trata de un factor que penaliza la distancia que separe los puntos muestrales de los puntos desconocidos y, por tanto, mientras mayor sea su magnitud, mayor penalización sufrirán las grandes distancias entre éstos, conllevado, por tanto, a una cada vez menor ponderación de su valor para el cómputo final. Por el contrario, si asignamos magnitudes muy bajas para este factor, obtendremos ponderaciones cada vez más significativas a pesar de las distancias que puedan separar a los puntos muestrales de los puntos desconocidos.

Atendiendo a los casos más extremos de estas casuísticas, en el caso de obtener un factor $p = 0$, encontraríamos que todas las ponderaciones ω_i serían iguales a 1 y que, por tanto, todas las observaciones muestrales son consideradas de forma equivalente, sin importar las distancias entre ellas y el punto a conocer. Así mismo, en el caso de obtener un factor “p” extremadamente elevado, únicamente se considerarían los pocos puntos más cercanos al punto desconocido, o incluso ningún punto muestral, para efectuar el cómputo de su valor. Normalmente, el factor “p” suele estar situado entre las magnitudes mayor que 0 y 3, siendo 2 el valor por defecto en la mayoría de los sistemas de cómputo geoestadístico.

En lo que corresponde a nuestro análisis, resulta útil obtener una representación de la concentración de los contaminantes principales (NO₂, O₃ y materias particuladas) en toda la

extensión del AYUNTAMIENTO DE MADRID. Para ello, antes de proceder con el cómputo de los valores desconocidos a través del método de la ponderación de la distancia inversa, es necesario abarcar las referencias correspondientes de cada punto contenido en nuestro territorio de análisis. Para ello, hemos decidido utilizar una cuadrícula que divida toda la extensión de nuestros puntos muestrales registrados sobre la región de estudio, y que dicha división genere una ingente cantidad de cuadros que se asemejen a los valores experimentados en cada punto posible del territorio a una resolución determinada. Nuevamente, sin embargo, resulto indispensable adaptar la cuadrícula, generada por librerías como *gstat* y *stars*, al sistema de referencias de coordenadas (CRS) y las proyecciones del mapa utilizado de conformidad con los estándares de ETRS89 / UTM zone 30, EPSG:25830.



Fig. 23 Cuadrícula de representación de ubicaciones del Ayuntamiento de Madrid con valores desconocidos de contaminación.

Por tanto, esta cuadrícula sirve para representar todas aquellas ubicaciones del AYUNTAMIENTO DE MADRID. Dentro de estos cuadros, encontraremos algunas estaciones de monitorización con datos conocidos, ya que ésta representa la totalidad del territorio. No obstante, y de manera más importante, esta cuadrícula también incluye todos los otros puntos de la provincia para los cuales

no tenemos valores de contaminación registrados y sobre los cuales podremos utilizar el método de la ponderación de la distancia inversa para estimarlos.

Una vez obtenida dicha cuadrícula para estimar la concentración sobre todo el área de la provincia de Madrid, procedemos a computar los niveles de contaminación estimados para los mismos. La librería de *gstat* permite realizar estas predicciones de manera sencilla, al introducir el conjunto de datos georreferenciados, en este caso los datos de las estaciones de monitorización con su latitud y longitud, y las ubicaciones de los puntos sobre los cuales se quiere realizar las estimaciones. A modo de obtener una mejor visualización, hemos decidido elaborar el gráfico a partir de las concentraciones de dióxido de nitrógeno registradas en cada estación de monitorización y las consiguientes estimaciones para toda la cuadrícula de la provincia, dado su presencia de registro en todo el territorio.

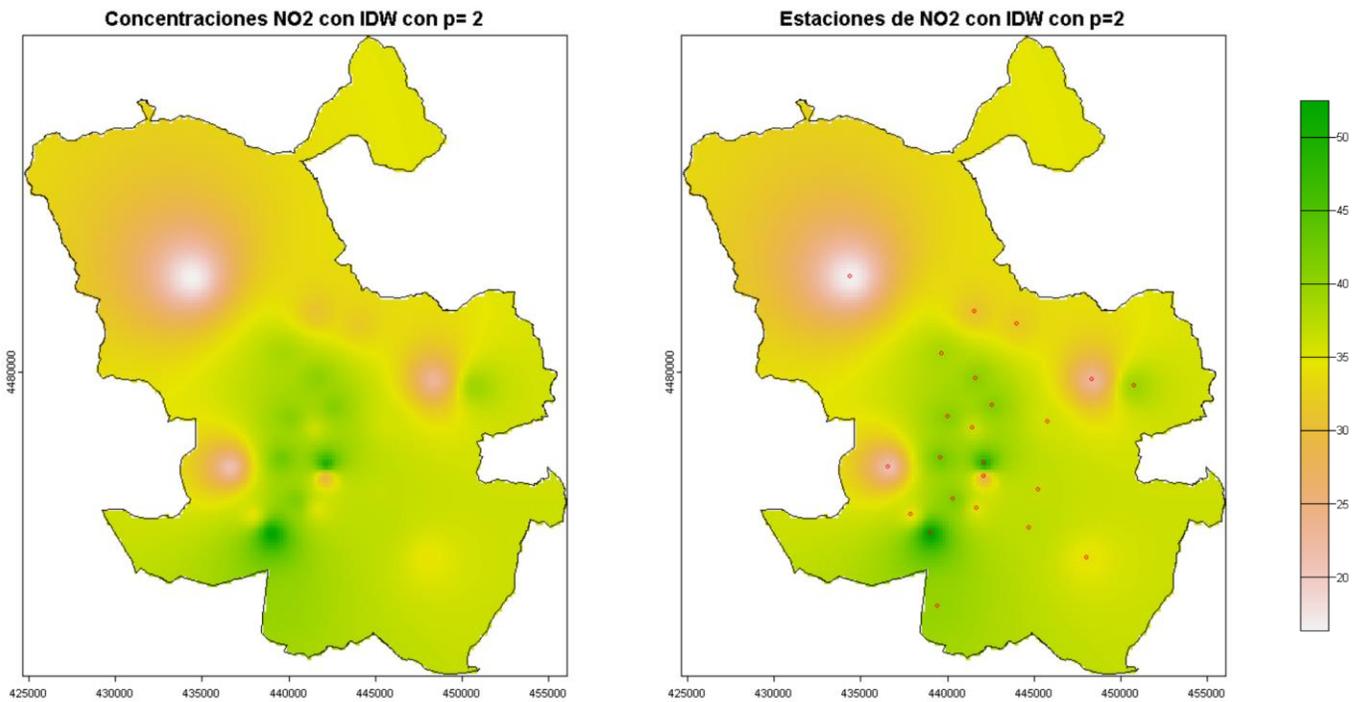


Fig. 24. Estimaciones de NO2 mediante método de IDW con factor "p" igual a 2.

Como podemos observar de la figura 24, se pueden confirmar las consideraciones expuestas previamente al analizar el marco teórico de este método de interpolación lineal. En primer lugar,

observamos grandes diferencias en cuanto a las concentraciones de NO₂ registradas en todo el territorio. Resulta claro que a medida que nos alejamos del centro urbano del Ayuntamiento de Madrid, por lo general, observaremos valores cada vez más bajos de NO₂. Muy probablemente, este fenómeno se ve explicado a su vez por las distribuciones vistas en las figuras 21 y 22, en donde observamos una sobrerrepresentación de los niveles de NO₂ en estaciones urbanas, así como en estaciones en donde el tráfico compone la principal fuente de contaminación en los alrededores de las estaciones de monitorización.

Ahora bien, y de forma más relevante para esta sección, podemos ver las características de la ponderación de la distancia inversa (IDW) como método de interpolación lineal. Al tratarse de un método determinístico y local, éste atribuye mayor importancia a las observaciones muestrales más cercanas a los puntos desconocidos para el cómputo de sus valores. En la figura 24 podemos ver como aquellas estaciones de monitorización rurales con valores bajos de concentración de NO₂ forman centros de gran importancia para todos los puntos desconocidos que los rodean, formando así los denominados “ojos de buey”, atribuibles a la falta de datos muestrales que contrasten esta elevada ponderación de puntos muestrales únicos. Así mismo, podemos ver en la parte baja de la misma figura que la ubicación de las estaciones de monitorización coinciden perfectamente con los epicentros de estos “ojos de buey”, y a medida que exista una mayor cantidad y variedad de valores muestrales en cercanía se obtienen predicciones mucho más suavizadas sobre los puntos desconocidos encontrados alrededor de estas estaciones de monitorización, como sucede a circunscribimos al área metropolitana de la ciudad.

Teniendo las estimaciones ya computadas, corresponde evaluar el poder de predicción de este método de la ponderación de la distancia inversa, es decir, el grado de acierto de estas estimaciones con lo esperado en la realidad. Para ello podemos recurrir a una medida de desempeño en particular como lo es la raíz del error cuadrático medio o RMSE (por sus siglas en inglés, *root mean square error*). Esta contrasta los valores predichos por un modelo con los valores actuales para una variable determinada, a través de diferencia (residuos). Una vez se obtiene esta diferencia, se elevan los residuos a la potencia de dos y se suman, dividiéndose luego por el número de residuos y se obtiene su raíz cuadrada. De esta forma, el grado de acierto se mide en la magnitud del RMSE, siendo el caso de que a menor RMSE, menor diferencias entre predicciones y valores observados.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Z(S_0)) - Z(S_i))^2}{N}}$$

Ahora bien, ya que no tenemos acceso a aquellos datos reales en puntos donde no hay estaciones monitorización, debemos emplear el método de validación cruzada para contrastar las estimaciones con datos de las estaciones de monitorización. Específicamente, hemos de utilizar el método k-folds de validación cruzada mediante el cual ignoramos aleatoriamente alguno de los datos utilizados para realizar las estimaciones y lo consideramos un punto desconocido a estimar, obteniendo así una predicción que puede o no coincidir con el valor observado. El proceso es repetido “k” veces, siendo k un número entre 1 y N.

Realizando dicha validación cruzada obtenemos un RMSE igual 9,41 aproximadamente. Dada la composición de esta medida de desempeño podemos asegurar que su valor se encuentra en las mismas unidades que la variable estimada y por tanto podemos decir que, en media, las estimaciones mediante IDW con un factor “p” igual a dos (2) se aleja de los valores de NO₂ observados por 9,41 µg/m³. Tomando en cuenta que la media de las concentraciones NO₂ estimadas se acerca a 36,22 µg/m³, esta distancia correspondería a un 26% del valor esperado de la variable.

Como hemos visto antes, este factor “p” se decide arbitrariamente. Por ello, resulta interesante determinar el valor de este factor a través de la minimización de una medida de desempeño como lo puede ser el RMSE. De esta manera utilizaríamos valores progresivos de “p”, en nuestro caso desde el 0,25 hasta el 3 en incrementos de 0,25, durante un nuevo proceso de validación cruzada para determinar cuál de estos valores minimiza el error en la predicción medido a través del RMSE. Aún más, podemos moderar el número de observaciones (vecinos) que utilizamos en cada estimación de un punto desconocido. Este número puede tomar los valore de 1 como mínimo y como máximo el número total de estaciones de monitorización, exceptuando el que se ignora en cada iteración de la validación cruzada.

Tras este proceso de optimización de los valores del factor “p” y el de número de vecinos, hemos sido capaces de identificar un valor de 0,25 para el factor “p” y un número de vecinos igual a 21. Con estas determinaciones, obtenemos una medida RMSE igual 8,63 µg/m³, aproximadamente.

Por tanto, se trata de una reducción de $0,78 \mu\text{g}/\text{m}^3$, mejorando nuestros estimados en un 8,3 % aproximadamente.

Método	Factor “p”	Nº Vecinos	Media de estimaciones ($\mu\text{g}/\text{m}^3$)	RMSE ($\mu\text{g}/\text{m}^3$)	RMSE/Media
IDW	2	25	36,224	9,411	25,98 %
IDW	0,25	21	36,224	8,629	23,82 %

A continuación, se puede ver, en la figura 26, el conjunto de estimaciones computadas por el

Fig. 25. Tabla comparativa de modelos de IDW elaborados.

método de la ponderación inversa utilizando los parámetros mencionados previamente. Podemos ver claramente como las apariciones de “ojos de buey” características de este método, se ven aminoradas en gran medida. Ello se debe a la disminución en el valor del factor “p”, y, por consiguiente, incrementando la importancia de los otros valores más lejanos sobre la estimación de todo punto desconocido. Este mismo efecto se puede ver en la zona más sur de la región de estudio en un aumento en el suavizado de las estimaciones, generando patrones amplios sobre estas zonas y pequeñas diferencias en diferentes direcciones.

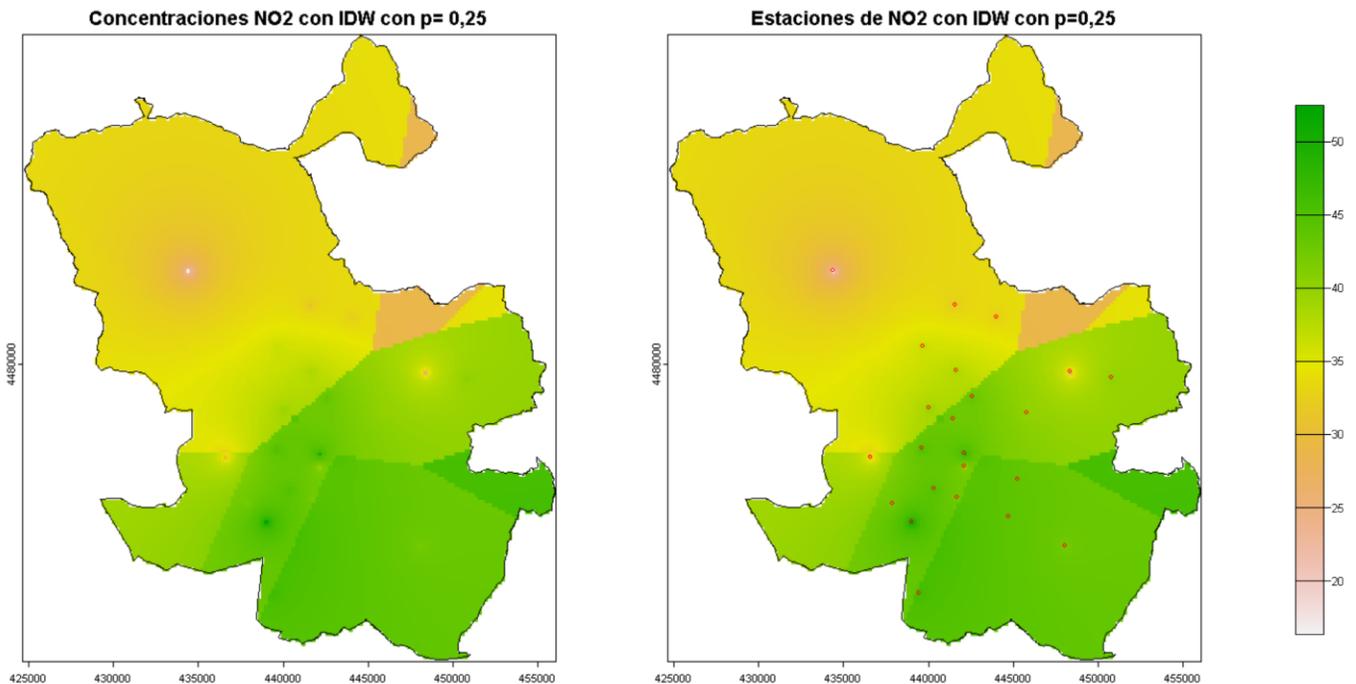


Fig. 26. Estimaciones de NO2 mediante método de IDW con factor “p” igual a 0,25.

3.2. Método Probabilístico Kriging

De forma similar a la sección anterior, hemos tenido la oportunidad de aproximarnos a las cuestiones esenciales del método de interpolación espacial kriging a lo largo de este trabajo. Así como el método de ponderación de la distancia inversa, el método kriging consiste en la estimación de valores desconocidos sobre puntos distintos a los ubicados en puntos muestrales con valores ya conocidos.

Así mismo, la función que describe dicho método de interpolación espacial se asemeja a la descrita para la ponderación de la distancia inversa (1). Se trata de un sumatorio que recorre todos los puntos muestrales N , empezando por el primero, y computando el producto entre los pesos de cada observación muestral (ω_i) y el valor de las mismas ($Z(S_i)$). La gran diferencia que separa a estos métodos de interpolación espacial se encuentra en la forma de determinar los pesos para cada observación muestral (ω_i).

A diferencia del método de la ponderación de la distancia inversa, el método kriging se trata de un estimador lineal insesgado que produce el menor error cuadrático medio en sus estimaciones. Para ello, el cómputo de las ponderaciones de cada valor muestral dos propiedades. En primer lugar, se debe garantizar la propiedad de insesgamiento, vista en la fórmula (3), mediante la cual la suma de las ponderaciones de todas las observaciones debe de ser igual a 1, evitando así la sobrerepresentación de algunos valores muestrales en el cómputo de las estimaciones (GIRALDO HENAO, Introducción a la Geoestadística).

$$E(Z^*(S_0)) = E(Z(S_0)) \quad (3)$$

Como podemos ver de la identidad (3), el método kriging constituye un método de interpolación espacial insesgado, en tanto en cuanto el valor esperado de las estimaciones sobre los valores de puntos desconocidos, $E(Z^*(S_0))$, sea igual al valor esperado de los verdaderos valores en dichos puntos, $E(Z(S_0))$. Ahora bien, esta igualdad también conlleva otra consideración importante. Desarrollando la ecuación, como se puede ver en los puntos (4), (5) y (6), se puede ver como al asumir proceso estacionario de media ' m ' desconocida, se arriba a la necesaria condición de que la suma de las ponderaciones ω_i sea igual a 1.

$$E\left(\sum_{i=1}^N \omega_i * Z(S_i)\right) = m \quad (4)$$

$$\sum_{i=1}^N \omega_i * E(Z(S_0)) = m \quad (5)$$

$$\sum_{i=1}^N \omega_i * m = m \quad (6)$$

Finamente, con una simple agrupación de los momentos descritos en la ecuación (6) podemos ver como en las ecuaciones descritas en el número (7) se resuelve dicha cuestión, arribando finalmente a la prueba de la necesidad de tener un sumatorio de las ponderaciones equivalente a 1.

$$m \sum_{i=1}^N \omega_i = m \rightarrow \sum_{i=1}^N \omega_i = 1 \quad (7)$$

Entendida esta cuestión, procedemos a analizar la forma de cómputo de, propiamente, de las ponderaciones ω_i atribuidas a cada observación muestral. Para ello, se parte de la consideración del método de kriging como uno que minimiza la varianza de predicción o la varianza de la diferencia entre el valor predicho y el valor actual, como estimador que produce el menor error medio de predicción (GIRALDO HENAO, Introducción a la Geoestadística). Esta condición se ve descrita en la ecuación (8) a continuación.

$$V[Z^*(S_0) - Z(S_0)] \quad (8)$$

El desarrollo de esta función conlleva la toma en cuenta de las covarianzas entre los puntos muestrales y los puntos a predecir. Así mismo, y como hemos expuesto previamente, la propiedad de insesgamiento condiciona dicha minimización de la varianza a respetar la necesidad de que la suma de las ponderaciones sea igual a 1. Todo ello resulta en un complejo sistema de ecuaciones contando con (n+1) ecuaciones y (n+1) incógnitas que matricialmente puede ser descrito como se ve a continuación (9) (GIRALDO HENAO, Introducción a la Geoestadística).

$$C_{ij} \cdot \omega = C_{i0} \rightarrow \omega = C_{ij}^{-1} \cdot C_{i0} \quad (9)$$

Por tanto, el cálculo de las ponderaciones para el método kriging corresponde con el producto matricial de la matriz inversa de las covarianzas entre datos muestrales (C_{ij}^{-1}) y de la matriz que contienen las covarianzas de todos los valores de los puntos muestrales con el punto cuyo valor debemos estimar (C_{i0}).

Ahora bien, esta relación matemática entre las matrices de covarianzas puede ser representada gráficamente a través de un semivariograma. Éste representa las covarianzas presentes entre pares

de datos muestrales observados a distintas distancias. Cada par viene determinado por la distancia que separa cada punto muestral del otro y se toman tantas combinaciones entre puntos muestrales como existan por cada unidad de distancia que se elija para determinar las covarianzas (COLUMBIA UNIVERISTY IRVING MEDICAL CENTER, Kriging Interpolation). En realidad, se trata de una función que determina las disimilitudes entre las observaciones a partir de las distancias que las separan (PYRCZ, 2018). De manera más precisa, el semivariograma representa un medio (1/2) del promedio de los cuadrados de las diferencias entre valores muestrales separados por una distancia “ h ”, visto en la función a continuación (semivarianza) (10).

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (Z(S_i) - Z(S_i + h))^2 \quad (10)$$

A partir de estas representaciones gráficas y de sus descripciones a través de funciones que provean el mejor encaje en términos de error cuadrático medio, podremos determinar las correlaciones y consiguientemente las covarianzas de los datos muestrales y por tanto poder obtener las ponderaciones correspondientes.

Atendiendo al caso concreto, debemos empezar mencionando que para la utilización del método kriging, hemos preferido circunscribirnos a los datos provenientes de las estaciones de monitorización comprendidas dentro del AYUNTAMIENTO DE MADRID. La razón de ello yace en las estructuras más adecuadas de datos que se encuentran disponibles en los valores registrados por estas estaciones de monitorización en comparación con las del resto de la COMUNIDAD DE MADRID.

Adicionalmente, y como tuvimos ocasión de analizar en otros apartados de este trabajo, las concentraciones de contaminantes como el NO₂ se encuentran en mayor abundancia en estaciones de monitorización de tipo de área urbana y aquellas expuestas al tráfico como fuente principal de contaminación.

En este sentido, resulta imperativo comenzar por la descripción gráfica de las covarianzas entre observaciones a lo largo del espacio muestral, es decir, por semivariograma de las concentraciones de NO₂ a las distintas distancias posibles, basado en las distintas ubicaciones de las estaciones de monitorización. En la siguiente figura, se puede ver, por una parte, los valores de semivarianza entre pares plasmados en función de las distancias que los separan, y, por otra, se puede ver la función que produce el mejor encaje con dichas observaciones. La función que en nuestro caso produce el mejor encaje es una función de tipo lineal, caracterizada por tener un crecimiento equiparado entre la semivarianza que experimentan todos los pares de valores y el crecimiento en las distancias que los separan (GIRALDO HENAO, Introducción a la Geoestadística). Los parámetros de esta confirman: una meseta (*sill*) de aproximadamente 80 de semivarianza, correspondiente con la varianza observada en los datos muestrales; un rango de aproximadamente 8.500 metros, indicando la distancia a partir de la cual no obtenemos información de la estructura espacial de la región de estudio; y un efecto pepita (*nugget*) de aproximadamente 50 de semivarianza, indicando la “discontinuidad puntual del semivariograma en el origen” (GIRALDO HENAO, Introducción a la Geoestadística, p. 24).

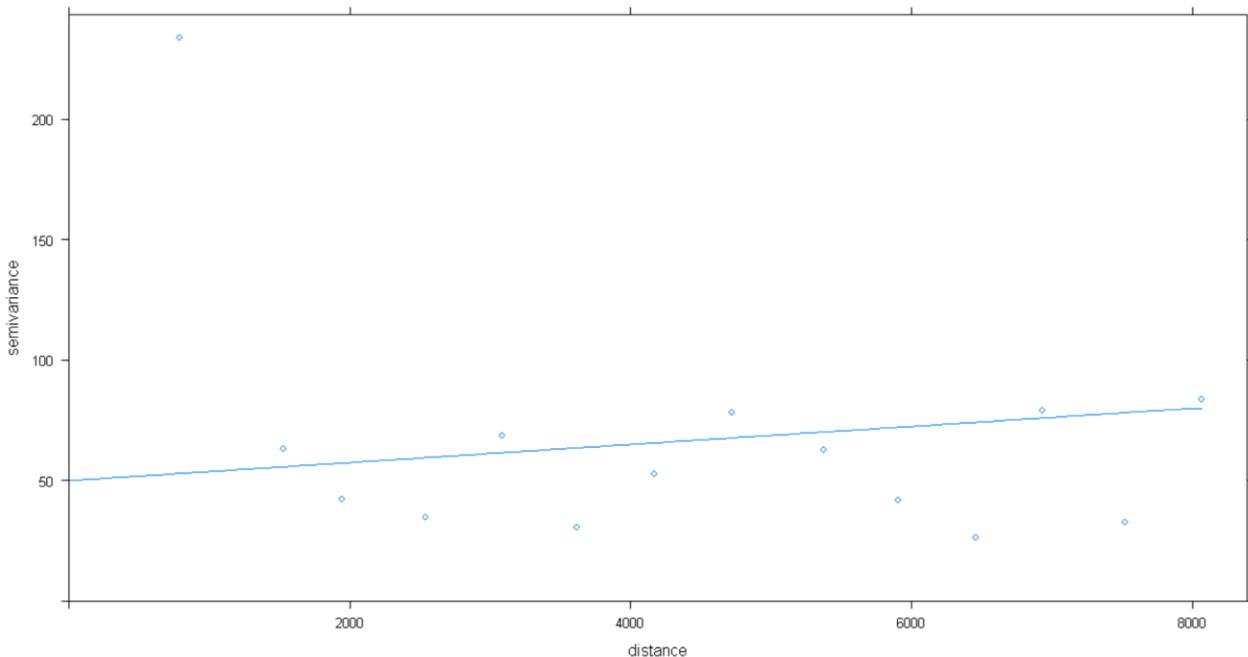


Fig. 27. Semivariograma y función de mejor encaje de tipo lineal.

Teniendo dicha función de la semivarianza, resulta simple proceder al cálculo de los pesos para las estimaciones correspondientes.

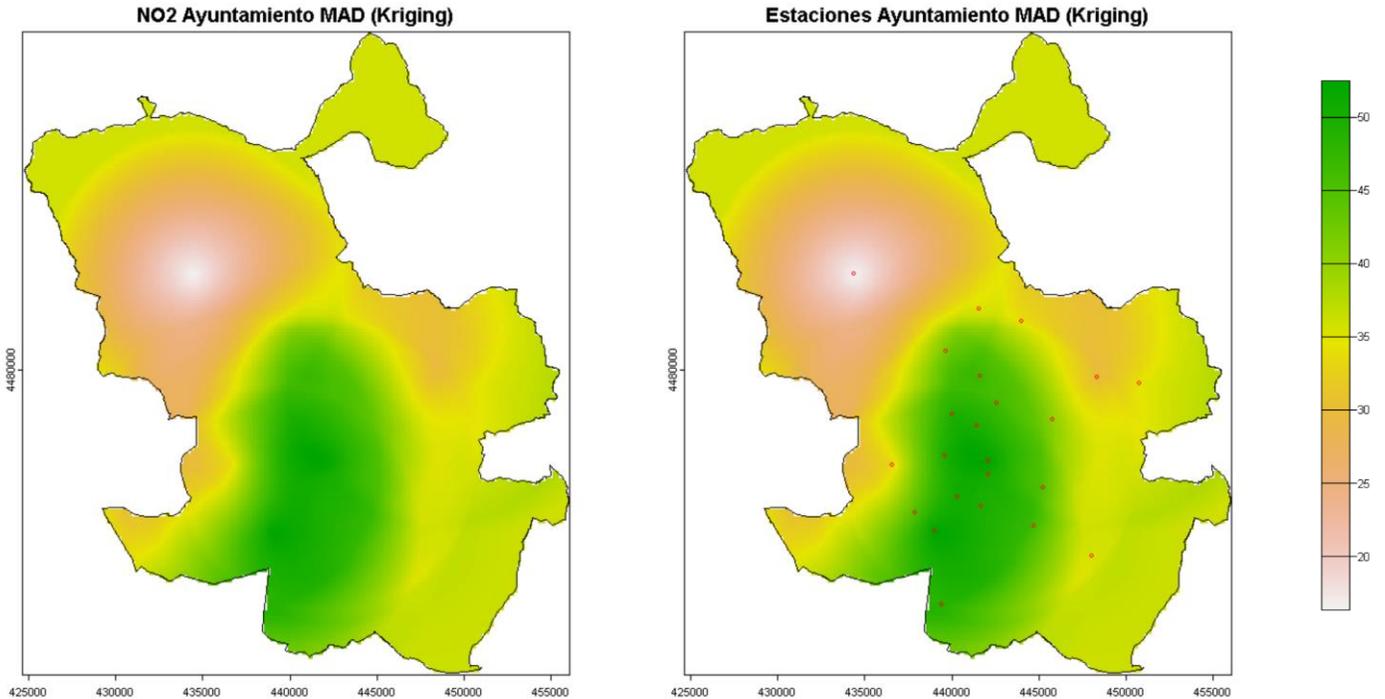


Fig. 28. Estimaciones de concentración de NO₂ en el Ayuntamiento de Madrid utilizando el método kriging.

Como podemos observar en la figura 28, resultan evidentes las diferencias que existen entre ambos interpoladores espaciales. A diferencia de la visualización elaborada a partir del método de la ponderación de la distancia inversa, esta visualización nos permite observar la presencia de dos grandes grupos diferenciados de valores de concentraciones de NO₂ en el AYUNTAMIENTO DE MADRID. Al parecer, existe una diferenciación en la dirección Norte-Sur que nos indica la aglomeración de altas concentraciones de NO₂ en el tramo centro-sur del AYUNTAMIENTO. Posiblemente, dicha diferenciación proceda de la tipología de área en las cuales se encuentran estas zonas de monitorización. Encontrados seguramente las de menor latitud en áreas de tipo urbano y probablemente mayormente expuestas a fuentes de contaminantes como el tráfico.

Por otra parte, podemos ver que a diferencia del método de IDW, no observamos la formación de “ojos de buey” producto de una sobre ponderación de datos muestrales locales y cercanos sobre aquellos desconocidos a sus alrededores. En cambio, se trata de una nube de estimaciones que

aparenta tener un suavizado incomparable con el que hemos podido ver en la fig. 26 sobre la misma área de cobertura en el AYUNTAMIENTO DE MADRID. Aún más, dadas las características de este interpolador espacial, resulta probable que algunos de los valores estimados en celdas de la cuadrícula coincidentes con las ubicaciones de estaciones de monitorización no correspondan con los valores actuales observados. No obstante, y como ya hemos podido comprobar, en media, este método de interpolación lineal tendrá un menor error de predicción.

Para determinar el desempeño de este modelo de interpolación espacial, utilizaremos el método de la validación cruzada para arribar a la medida RMSE, así como hicimos con el método IDW. A continuación, mostramos una tabla comparativa de todos los modelos evaluados. En ella, podemos ver como la consecuencia del efecto suavizado se manifiesta en una media de estimaciones aminoradas, ya que reduce los valores extremos por las cuales estima la variable. Así mismo, este efecto queda evidenciado en la reducida varianza sobre las estimaciones generadas, y muestra el poder de suavizado que tienen un factor “p” cercano a 0, como lo es el óptimo en el método IDW en términos de RMSE. Por último, podemos ver un RMSE menor comparado con las implementaciones IDW en cualquier versión, en términos absolutos, no obstante, en términos relativos se muestra mejor predicción el método IDW con los parámetros óptimos descritos.

Método	Factor “p”	Nº Vecinos	Función Semivariograma	Media de estimaciones ($\mu\text{g}/\text{m}^3$)	Varianza de estimaciones	RMSE ($\mu\text{g}/\text{m}^3$)	RMSE/Media
IDW	2	25	NA	36,224	15,828	9,411	25,98 %
IDW	0,25	21	NA	36,224	0,552	8,629	23,82 %
Kriging Ordinario	NA	25	Lineal	33,740	5,228	8,563	25,37 %

Fig. 29. Tabla comparativa de modelos de interpolación espacial elaborados.

Finalmente, resulta interesante ver la elaboración de la fig. 30 en donde se detalla la varianza que se tiene sobre cada estimación computada por el método kriging. Aquí, podemos observar el nivel de incertidumbre que hay sobre las predicciones de todo el territorio del AYUNTAMIENTO. No obstante, observamos que a medida que nos acercamos al conjunto de estaciones de monitorización

aglomeradas, es decir, cerca del área metropolitana de Madrid, la varianza sobre las estimaciones disminuye gradualmente, encontrándose aquí los menores márgenes de error en las predicciones. Las zonas más remotas del AYUNTAMIENTO, aquellas en donde no hay una gran densidad poblacional, se ven más sujetas a un error en la predicción mediante este método con la actual configuración de la red de estaciones y su distribución a lo largo del territorio.

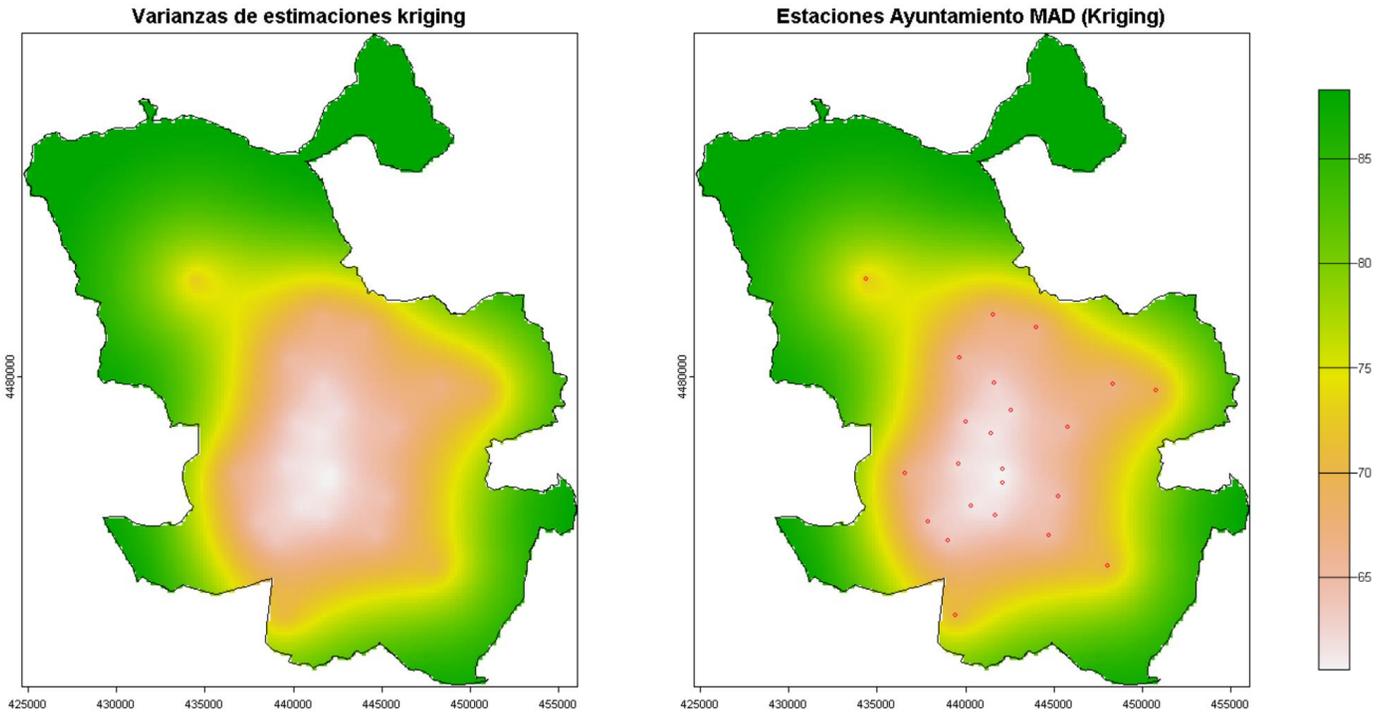


Fig. 30. Estimaciones de las varianzas de las estimaciones de NO₂ en el Ayuntamiento de Madrid utilizando el método kriging.

CAPÍTULO IV. DISTRIBUCIÓN ÓPTIMA DE LAS ESTACIONES DE MONITORIZACIÓN DE LA RED DE VIGILANCIA DEL AYUNTAMIENTO DE MADRID.

Tras el análisis exploratorio de la composición de la red de vigilancia de calidad del aire del AYUNTAMIENTO DE MADRID, así como de la aproximación a métodos de interpolación espacial que nos permiten conocer los posibles valores de contaminación lo largo de todo el territorio, no sabría preguntar si el diseño y gestión de esta red resulta óptimo para el conocimiento de la calidad del aire a cada instante. De manera más precisa, cabría ponderar las posibilidades de aumentar la capacidad predictiva de sistema de vigilancia de alguna forma u otra.

Ante esta duda, resulta inmediato pensar en la implantación de estaciones de monitorización en aquellas zonas que actualmente no se vean dotadas de las mismas en una proporción similar a la

que abarca el registro de calidad del aire del área metropolitana de Madrid. Ello resultaría en un mayor número de datos muestrales sobre los cuales basarse métodos de interpolación espacial, como los estudiados en la sección anterior, e introduciendo datos reales con los cuales contrastar predicciones anteriores. Así mismo, se podría pensar en actualizar los sistemas de monitorización ubicados dentro de cada estación. La introducción de nueva tecnología en la forma de analizadores individuales, seguramente, pueda resultar en un aumento de la precisión de las concentraciones registradas, así como en el incremento del área monitorizada y de la reducción de posibles datos perdidos.

Ahora bien, las alternativas mencionadas anteriormente, a pesar de producir mejoras instantáneas en la capacidad predictiva del sistema de vigilancia en cuestión, conllevan un considerable coste económico que puede resultar problemático a la hora de implementar dichas mejoras. Por ello, en esta sección nos dedicamos a evaluar una opción que conlleva un coste económico inmensamente menor y con un potencial de mejora predictiva significativo; la redistribución de las actuales estaciones de monitorización a lo largo del territorio del AYUNTAMIENTO DE MADRID.

Conociendo la actual ubicación de las estaciones de monitorización, la estructura de dependencia espacial que se recoge en la región de estudio (semivariograma) y las estimaciones y varianzas de las mismas que resultan de métodos como el kriging, resultará, relativamente, sencillo encontrar la mejor distribución de las estaciones actuales para mejorar la capacidad predictiva de la red de vigilancia en su conjunto.

Se trata, por tanto, de un proceso de optimización, y de forma más precisa, se trata de una minimización de la varianza sobre las estimaciones realizadas para niveles de contaminación en toda la región de estudio. En otras palabras, y de manera visual, se podría entender como la extensión de varianzas pequeñas como la vista sobre el área metropolitana de Madrid en la figura 30, a lo largo de la mayor cantidad de puntos de la región de todo el AYUNTAMIENTO.

Ahora bien, esta “*minimización de la varianza*”, puede implicar dos consecuencias distintas. Por un lado, se podría minimizar la varianza media sobre todas las estimaciones que cubren la región de estudio (11). Por otro lado, se podría minimizar la máxima varianza experimentada considerando todas las estimaciones de la región de estudio (12). La primera consecuencia derivaría en una distribución de las estaciones de monitorización, de forma tal que toda la región de estudio quede cubierta de la forma más homogénea posible. La segunda consecuencia, por el

contrario, conllevaría a una cobertura más heterogénea del área de estudio, pero aseguraría que ningún punto de ella este cubierto de forma pobre por la capacidad predictiva, y, por tanto, resulta idóneo cuando toda la región debe encontrarse bajo un umbral determinado, como resulta en nuestro caso sobre la calidad del aire y su regulación española y europea (ROSSITER & SRIVASTAVA, 2021).

$$\text{MIN}_i \frac{1}{N} \sum_{i=1}^N \widehat{\sigma}_i^2 \quad (11)$$

$$\text{MIN}_i \left(\text{MAX}_i \widehat{\sigma}_i^2 \right) \quad (12)$$

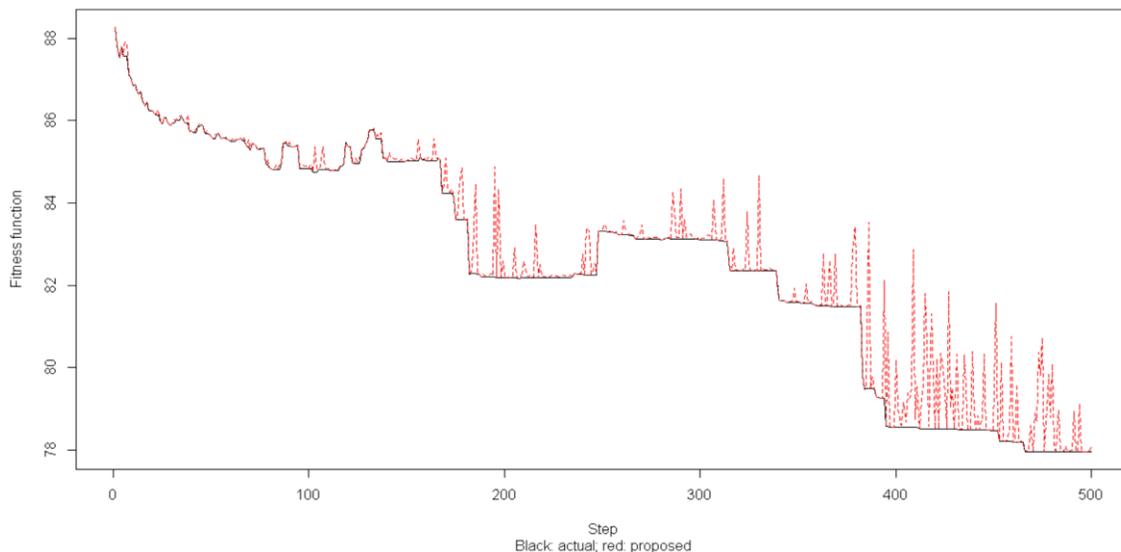
A la hora de aproximarnos a resolver este problema, podríamos pensar en aplicar la gran capacidad computacional que nos ofrece la tecnología y, ante ello, comparar las varianzas experimentadas para todas las ubicaciones posibles a lo largo de la región de estudio (ROSSITER & SRIVASTAVA, 2021). Sin embargo, una vista rápida a la sección donde hemos creado el cuadrículado que recoge el área de estudio en la figura 23 nos hará recordar el ingente número de posibles ubicaciones de muestreo, generando un proceso casi interminable considerando que, además, nos encontramos trabajando con 25 estaciones de monitorización que distribuir a lo largo de estas posibles ubicaciones.

Descartando rápidamente este método de trabajo, nos centraremos en el algoritmo de recocido simulado espacial (SSA, por sus siglas en inglés, *spatial simulated annealing*) como método de optimización para encontrar una solución a esta problemática. Éste se basa en la modificación repetitiva de la ubicación de una red de puntos muestrales, aleatoriamente distribuidos o fijados arbitrariamente, y basando el criterio de su movimiento en una reducción de la varianza, ya sea la varianza media o la máxima varianza, como se ha diferenciado previamente (ROSSITER & SRIVASTAVA, 2021). Es decir, este método toma una determinada distribución georreferenciada (ubicaciones) y, de forma aleatoria, elige algún punto muestral registrado y propone cambios de ubicación para el mismo, tomando en cuenta la estructura espacial (semivariograma). Esta propuesta es entonces aceptada o rechazada en función de su aminoración o no de la varianza de la región de estudio. Este proceso necesita de la fijación un número de iteraciones sobre la cual repetir el proceso y el cual se puede decidir arbitrariamente o fijar hasta encontrar una varianza deseada si es posible.

En nuestro entorno de programación (R), encontramos una función determinada que permite realizar este proceso de manera muy sencilla. Llamada *ssa.ok* (*spatial sample annealing. ordinary kriging*), y desarrollada por los investigadores ROSSITER & SRIVASTAVA, en su publicación del 2021, esta función únicamente necesita de: un diseño (distribución espacial) inicial de los puntos muestrales, en nuestro caso, las coordenadas de las estaciones de monitorización existentes; una serie de polígonos que describan la región de estudio, en nuestro caso los datos shapefile del AYUNTAMIENTO DE MADRID; un modelo de semivariograma que describa la estructura espacial de los datos y de la región de estudio, en nuestro caso, la elaborada en la sección anterior que sigue un modelo lineal; el número de interacciones deseadas para la computación, 500 en nuestro caso; y el criterio de optimización, ya sea la minimización de la varianza media o la máxima varianza.

Teniendo la posibilidad de determinar el diseño óptimo de la red de vigilancia de la región, hemos decidido computar este optimizando la minimización de la máxima varianza y realizar un análisis comparativo de la misma con el actual diseño y su distribución de varianza. No obstante, antes de ello, resulta interesante detallar el recorrido que realiza el algoritmo de recocido simulado espacial en búsqueda de su respectivo criterio de optimización. En la figura 31 podemos ver como el gráfico elaborado corresponde con el recorrido de dicho método en la minimización de la máxima varianza. Así mismo, se puede ver el recorrido verdadero del algoritmo en la línea negra y todos los cambios de ubicación propuestos pero rechazados por el aumento en la varianza, en su respectivo criterio de optimización particular.

Fig. 31. Recorrido de algoritmo de recocido simulado espacial en 500 iteraciones.



Así mismo, a continuación, se puede ver la figura 32, la cual señala las ubicaciones a lo largo de la región del AYUNTAMIENTO DE MADRID, que se determinan óptimas candidatas para las actuales estaciones de monitorización con el objetivo de minimizar la máxima varianza registrada en alguna estimación sobre toda la región de estudio. Como se puede ver, en claro contraste con la figura 32, no existe un aglomerado de estaciones con tal cercanía como la que describe la actual configuración de la red en el área metropolitana de Madrid. En cambio, se trata de una distribución bastante equitativa del número finito de estaciones con el que estamos trabajando. De forma similar, se puede ver una extendida aminoración de las varianzas recogidas sobre las estimaciones de todo el territorio, estando claramente resaltado por un color naranja, similar al que se extiende por el área metropolitana de Madrid en la figura 30.

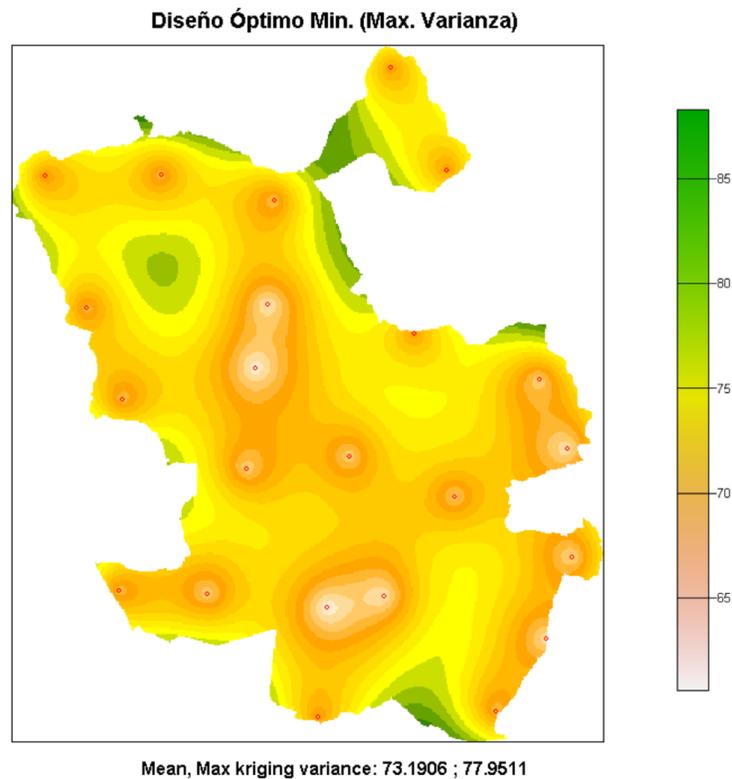


Fig. 32. Varianzas de estimaciones de NO₂ en el Ayuntamiento de Madrid utilizando el método kriging sobre el diseño óptimo propuesto.

Finalmente, resulta imperativo observar los datos de cada uno de los diseños propuestos, el óptimo descrito arriba y el actual visto previamente, para determinar así las diferencias entre ambos.

Medida	Diseño Actual		Diseño Óptimo	
	Estimaciones NO2 ($\mu\text{g}/\text{m}^3$)	σ^2 de estimaciones NO2 ($(\mu\text{g}/\text{m}^3)^2$)	Estimaciones NO2 ($\mu\text{g}/\text{m}^3$)	σ^2 de estimaciones NO2 ($(\mu\text{g}/\text{m}^3)^2$)
Media	33,74	76,72	36,13	73,19
Mediana	33,74	77,12	36,12	73,24
Máximo	38,96	88,28	43,19	77,95
Mínimo	27,63	60,59	28,78	68,18
Cuartil 1 (0.25)	32,38	69,68	35,37	72,22
Cuartil 3 (0.75)	34,84	84,21	36,93	74,23

Fig. 33. Tabla resumen comparativa de los diseños de la red de vigilancia del Ayuntamiento de Madrid.

Para ello, podemos ver que en la figura 33, expuesta arriba, se recogen los datos estadísticos principales sobre las estimaciones y varianzas de las mismas, por cada diseño de la red de vigilancia. En primer lugar, se observa una efectiva mejora sobre el valor de la máxima varianza registrada, encontrando nuestro diseño óptimo una minoración significativa en 10,33 unidades de varianza, equivalentes a unas 3,21 unidades $\mu\text{g}/\text{m}^3$. Además de ello, el nuevo diseño de la red de vigilancia disminuye tanto la media como la mediana de la varianza estimada a lo largo de la región de estudio, por lo cual asegura una mejor cobertura sobre la calidad del aire de toda la región considerada. Por último, sin embargo, es importante mencionar, que el diseño óptimo que proponemos no ofrece tanta especificidad sobre la varianza en puntos concretos sobre el área metropolitana de Madrid, como lo hace el diseño actual. Debido a la aglomeración de estaciones de medición en esta área, el diseño actual es capaz de ofrecer una varianza menor sobre esta área tan poblada, correspondiente con el mínimo visto ($60,59 (\mu\text{g}/\text{m}^3)^2$).

CAPÍTULO V. CONCLUSIONES

Luego de haber realizado este extenso análisis del estado actual de la monitorización de la calidad del aire a distintos niveles, así como de las cuestiones propias de los distintos métodos de interpolación espacial, podemos establecer las siguientes conclusiones:

Primera. El acceso a un aire limpio es reconocido como un derecho fundamental por parte de las instituciones europeas y los estados miembros de la Unión. Sin embargo, se han violentado los umbrales máximos de calidad establecidos tanto por la normativa europea como por la nacional a lo largo de los años.

Segunda. Los contaminantes principales ubicados en territorio español son las partículas en suspensión (PM10 y PM2,5), el dióxido de nitrógeno (NO2) y el ozono troposférico (O3). Cada uno de ellos causa graves perjuicios en los sistemas respiratorios y cardiovasculares de los seres humanos, además de introducir fenómenos meteorológicos que destruyen nuestra vegetación cada día.

Tercera. Con el objetivo de obtener una imagen fidedigna de la calidad del aire, tanto a niveles históricos como actuales, los distintos entes gubernamentales han implementado iniciativas que comprenden redes de vigilancia o monitorización basadas en estaciones permanentes de registro distribuidas a lo largo de los territorios en cuestión. No obstante, y a pesar de sus potentes capacidades de registro, dichos sistemas como el ICA, CAMS, entre otros, muestran áreas de mejora en la estimación de la calidad del aire para todo el territorio abarcado en lugar de su mera constatación en lugares puntuales del mismo, coincidentes con las estaciones de monitorización. Así mismo, existen una cantidad de iniciativas privadas, como BREEZOMETER, cuyo objetivo viene siendo la estimación de los niveles de contaminación a resoluciones inigualables, aportando un nivel de granularidad sobre la calidad del aire de suma precisión.

Cuarta. Existe una ingente cantidad de métodos que permiten conocer valores espaciales en puntos desconocidos a través de observaciones puntuales distribuidas a lo largo de la misma región de espacio. Dentro de los más conocidos encontramos el método de la ponderación de la distancia inversa y el método kriging, los cuales se diferencian principalmente en la importancia relativa de cada observación muestral para el cómputo de las estimaciones respectivas, estando el primero de estos determinado por la distancia que separa a dichos puntos, principalmente, y estando el

segundo determinado por la estructura espacial de la variable a estimar a lo largo de la región de estudio.

Quinta. Dentro de los métodos de interpolación espacial que ofrecen una posible solución a esta problemática, el método de kriging ordinario se erige como mejor estimador de niveles de concentración de NO₂ en el AYUNTAMIENTO DE MADRID. No obstante, mediante la optimización de los parámetros “p” y número de vecinos utilizados por el método de la ponderación de la distancia inversa, se pueden obtener mejores estimaciones que las computadas mediante el método kriging, aplicando valores tan bajos como 0,25 para el factor “p” de manera que se produzca un suavizado que reduzca el RMSE de dicho método.

Sexta. La distribución de las concentraciones estimadas para el NO₂ en el AYUNTAMIENTO DE MADRID se muestra coincidente para ambos métodos de interpolación espacial. Existe una clara aminoración de la concentración de dicho contaminante en las zonas noroestes de la región, y encontrando valores muy elevados a lo largo del centro, coincidente con el área metropolitana de Madrid, y extendiéndose en las direcciones suroeste de la región.

Séptima. Dichos resultados confirman la relación que existe entre las fuentes de contaminantes como las fuentes de actividad humana y la concentración de NO₂ en la región del AYUNTAMIENTO DE MADRID. Así mismo, las concentraciones de NO₂ resultan más elevadas en estaciones ubicadas en áreas urbanas en comparación con áreas de tipo suburbano o rural.

Octava. Finalmente, existe un margen de mejora en cuanto a la distribución de las actuales estaciones de monitorización extendidas a lo largo del AYUNTAMIENTO DE MADRID. Una redistribución coincidente con las ubicaciones previstas en el último capítulo de este trabajo reduciría la máxima varianza prevista para toda estimación de los niveles de NO₂ en la región. Así mismo, dicha configuración reduciría la varianza media de las estimaciones de NO₂ sobre toda la región considerada. No obstante, dada su desconcentración sobre el área metropolitana de Madrid, se perdería la precisión obtenida sobre esta sección tan densamente poblada.

BIBLIOGRAFÍA

ArcGIS Pro (s.f.). Cómo funciona la interpolación de distancia inversa ponderada.

<https://pro.arcgis.com/es/pro-app/latest/help/analysis/geostatistical-analyst/how-inverse->

[distance-weighted-interpolation-works.htm#:~:text=Como%20se%20ha%20mencionado%20anteriormente,distancia%2C%20las%20ponderaciones%20disminuyen%20r%C3%A1pidamente.](#)

Ayuntamiento de Madrid (s.f.). *Bases de datos y publicaciones. Calidad del Aire. En tiempo real.* <https://airedemadrid.madrid.es/portales/calidadaire/es/Bases-de-datos-y-publicaciones/Bases-de-datos-de-calidad-del-aire/En-tiempo-real/?vgnextfmt=default&vgnextchannel=650a89e859517710VgnVCM1000001d4a900aRCRD>

Bhandari, S. *et al.* (2022). Spatial and Temporal Analysis of Impacts of Hurricane Florence on Criteria Air Pollutants and Air Toxics in Eastern North Carolina. *Internal Journal of Environmental Research and Public Health*, 19, 1757. <https://doi.org/10.3390/ijerph19031757>

BreezoMeter (s.f.) *Air quality map. Real-time & Street-level Air Quality Information in Madrid.* <https://www.breezometer.com/air-quality-map/air-quality/spain/madrid>

BreezoMeter & Google (s.f.) *The Ultimate Guide to BreezoMeter's Air Quality Technology.* https://www.breezometer.com/pdfs/Ultimate-Guide-to-BreezoMeter_s-Air-Quality-Technology.pdf

Cañada Torrecilla, R. *et al.* (2010). Interpolación espacial y visualización cartográfica para el análisis de la justicia ambiental: ensayo metodológico sobre la contaminación por partículas atmosféricas en Madrid. En Ojeda, J., Pita, M.F. y Vallejo, I. (Eds.), *Tecnologías de la Información Geográfica: La Información Geográfica al servicio de los ciudadanos* (pp. 691-715). Secretariado de Publicaciones de la Universidad de Sevilla. <https://idus.us.es/handle/11441/66612>

CAMS (s.f.) *What we do. Observations.* <https://atmosphere.copernicus.eu/observations>

Comunidad de Madrid (s.f.) *Calidad del aire y salud.* <https://www.comunidad.madrid/servicios/salud/calidad-aire-salud#ozono-o3>

Comunidad de Madrid (s.f.). *Red de Calidad del Aire de la Comunidad de Madrid. Índice de Calidad del Aire.* http://gestionamadrid.org/azul_internet/run/j/AvisosAccion.icm

Columbia University Irving Medical Center (s.f.). Kriging Interpolation. Mailman School of Public Health. <https://www.publichealth.columbia.edu/research/population-health-methods/kriging-interpolation>

Copernicus (s.f.). *Sobre Copernicus*. <https://www.copernicus.eu/es/sobre-copernicus>

Copernicus (s.f.). *European air quality. Air quality map*. <https://www.regional.atmosphere.copernicus.eu/>

Ecologistas en Acción (2022). *La calidad del aire en el Estado español durante 2021*. <https://www.ecologistasenaccion.org/wp-content/uploads/2022/06/informe-calidad-aire-2021.pdf>

Fernández Rivas, L. (2015). *El papel de los óxidos de nitrógeno en el Cambio Climático. Efectos sobre la salud*. [Trabajo de Fin de Grado]. Universidad Complutense. [https://eprints.ucm.es/id/eprint/48627/1/LAURA%20FERNANDEZ%20RIVAS%20\(1\).pdf](https://eprints.ucm.es/id/eprint/48627/1/LAURA%20FERNANDEZ%20RIVAS%20(1).pdf)

Generelitat de Catalunya (s.f.). *El ozono troposférico en Catalunya. ¿Qué es el ozono troposférico?* Medio Ambiente y Sostenibilidad. https://mediambient.gencat.cat/es/05_ambits_dactuacio/atmosfera/qualitat_de_laيرة/avaluacio/campanya_ozo/que_es_lozo_troposferic/

Giraldo Henao, R. (2002). *Introducción a la Geoestadística. Teoría y Aplicación*. Universidad Nacional de Colombia. https://geoinnova.org/wp-content/uploads/2021/08/LIBRO_-DE-_GEOESTADISTICA-R-Giraldo.pdf

Gobierno de la Rioja (s.f.). Calidad del Aire y Cambio Climático. Óxidos de nitrógeno. Larioja.org. <https://www.larioja.org/medio-ambiente/es/calidad-aire-cambio-climatico/calidad-aire/evolucion-principales-contaminantes/oxidos-nitrogenos#:~:text=Los%20%C3%B3xidos%20de%20nitr%C3%B3geno%20de,de%20olor%20asfixiante%20y%20t%C3%B3xico.>

Lobelia Air (s.f.) *Calidad del aire en Madrid*. <https://aire-madrid.lobelia.earth/ayuntamiento-madrid/es/?species=PM10>

- Ministerio para la Transición Ecológica y el Reto Demográfico (s.f.). Calidad del aire. Visor de Calidad del Aire. <https://www.miteco.gob.es/es/calidad-y-evaluacion-ambiental/temas/atmosfera-y-calidad-del-aire/calidad-del-aire/visor/>
- Moral García, F.J. (2004). Aplicación de la geoestadística en las ciencias ambientales. Ecosistemas. Revista Científica y Técnica de ecología y Medio Ambiente, 13, (1) (pp. 78-86). <https://revistaecosistemas.net/index.php/ecosistemas/article/view/582>
- Orden TEC/351/2019, de 18 de marzo, por la que se aprueba el Índice Nacional de Calidad del Aire. *Boletín Oficial del Estado*, 28 de marzo de 2019. <https://www.boe.es/buscar/doc.php?id=BOE-A-2019-4494>
- Organización Mundial de la Salud (2021). *Directrices mundiales de la OMS sobre la calidad del aire*. https://www.miteco.gob.es/es/calidad-y-evaluacion-ambiental/temas/atmosfera-y-calidad-del-aire/guiaoms2021-spa_tcm30-530942.pdf
- Organización Mundial de la Salud (19 de diciembre 2022). *Contaminación del aire ambiente (exterior)*. [https://www.who.int/es/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](https://www.who.int/es/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health)
- Organización Panamericana de la Salud (2018). Contaminación del aire ambiental exterior y en la vivienda: Preguntas frecuentes. <https://www.paho.org/es/temas/calidad-aire-salud/contaminacion-aire-ambiental-exterior-vivienda-preguntas-frecuentes>
- Pyrcz, M.J. [GeostatsGuy Lectures] (23 de octubre de 2018). *10d data Analytics: Variogram Caculation*. [Archivo de Video]. YouTube. <https://www.youtube.com/watch?v=mzPLicovE7Q&t=834s>
- QGIS (s.f). *Manual de entrenamiento. Sistema de Referencia de Coordenadas*. https://docs.qgis.org/3.28/es/docs/gentle_gis_introduction/coordinate_reference_systems.html
- Quesada Muelas, L. (2019) *¿Qué es la interpolación espacial de datos?* Geoinnova. <https://geoinnova.org/blog-territorio/que-es-la-interpolacion-espacial-de-datos/>

Real Decreto 102/2011, de 28 de enero, relativo a la mejora de la calidad del aire. *Boletín Oficial del Estado*, 29 de enero de 2011. <https://www.boe.es/buscar/doc.php?id=BOE-A-2011-1645>

Rebollo, R. (2010). *Análisis de Señales en Geofísica*. Universidad Nacional de la Plata. <http://carina.fcaglp.unlp.edu.ar/senales/apuntes/Geoestadistica.pdf>

Resolución A/RES/76/300. Promoción y protección de los derechos humanos: cuestiones de derechos humanos, incluidos otros medios de mejorar el goce efectivo de los derechos humanos y las libertades fundamentales. *United Nations Digital Library*, de 28 de julio de 2022. <https://digitallibrary.un.org/record/3982508?ln=en>

Resolución de 2 de septiembre de 2020, la Dirección General de Calidad y Evaluación Ambiental, por la que se modifica el Anexo de la Orden TEC/351/2019, de 18 de marzo, por la que se aprueba el Índice Nacional de Calidad del Aire. *Boletín Oficial del Estado*, de 10 de septiembre de 2020. https://www.miteco.gob.es/es/calidad-y-evaluacion-ambiental/temas/atmosfera-y-calidad-del-aire/resolucion_02092020_modificacion_ica_tcm30-511596.pdf

Rossiter, D.G. y Srivastava, V. (2021). *Applied geostatistics. Exercise: Spatial simulated annealing*. University of Twente. https://www.css.cornell.edu/faculty/dgr2/_static/files/R_PDF/exSSA.pdf

Subdirección General de Aire Limpio y Sostenibilidad Industrial (2023). Tendencias de la Calidad del Aire en España 2001-2021. Secretaría General Técnica. Centro de Publicaciones. https://www.miteco.gob.es/es/calidad-y-evaluacion-ambiental/temas/atmosfera-y-calidad-del-aire/analisisdetendenciasdelosprincipalescontaminantesatmosfericos_tcm30-561228.pdf